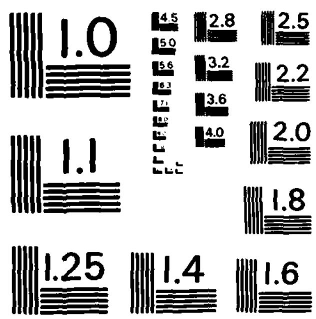
AD-A147 667   A TIME SERIES ANALYSIS OF RECOVERABLE SPARES REQUIREMENTS(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF SYSTEMS AND LOGISTICS T G LOCKETTE   1/1

UNCLASSIFIED   SEP 84 AFIT/GLM/LSM/84S-38   F/G 15/5   NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A147 667

A TIME SERIES ANALYSIS OF

RECOVERABLE SPARES REQUIREMENTS

THESIS

THOMAS G. LOCKETTE
GM-13

AFIT/GLM/LSM/84S-38

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

84 11 14 098

AFIT/GLM/LSM/84

A TIME SERIES ANALYSIS OF

RECOVERABLE SPARES REQUIREMENTS

THESIS

THOMAS G. LOCKETTE
GM-13

AFIT/GLM/LSM/84S-38

The contents of the document are technically accurate, and no sensitive items, detrimental ideas, or deleterious information are contained therein. Furthermore, the views expressed in the document are those of the author(s) and do not necessarily reflect the views of the School of Systems and Logistics, the Air University, the United States Air Force, or the Department of Defense.

Table of Contents

Page

AFIT/GLM/LSM/84S-38

A TIME SERIES ANALYSIS OF RECOVERABLE SPARES REQUIREMENTS

THESIS

Presented to the Faculty of the School of Systems and Logistics

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Logistics Management

THOMAS G. LOCKETTE, MBA

GM-13

September 1984

Preface

The purpose of this study was to accomplish a times series analysis of C-130 recoverable spares requirements to identify relationships within and between C-130 flying hours and spares requirements. The need for a better understanding of these relationships has been highlighted by recent highly publicized understatements of spares requirements. This study is not intended to develop a forecasting model, but rather, to demonstrate a methodology to be used to identify relationships within and between various logistics time series not only with C-130 aircraft but also with other weapon systems.

Although many have contributed to the accomplishment of this thesis, a special debt is owed to Major Carlos Talbott, my thesis advisor, who was so generous with his experience and expertise. I also thank Carla Siefert, who typed this thesis, for her help in producing a quality final product. Finally, I owe my wife Kathy a special thanks for her patience, understanding, and support during this effort.

Thomas G. Lockette

List of Figures

AFIT/LSM/84S-38

Abstract

The need for a better understanding of relationships between various logistics variables and spares requirements has become evident due to recent understatements of recoverable spares requirements. This thesis used Box-Jenkins time series analysis techniques to identify the linear relationships within and between C-130 spares requirements and flying hours. A methodology will be demonstrated which could then be applied to other variables (fleet value, fleet age, etc.) and other aircraft. A more accurate forecasting technique, than the linear regression technique presently used, should then evolve as more relationships are discovered. The TIMES package on the AFIT Harris computer was used to accomplish the analysis.

The analysis identified relationships within and between the two time series which are not currently considered in the linear regression forecasting model. A model which recognizes these additional relationships (time lags, seasonality, etc.) provides a better "fit" to the historical data, mesured by a chi-square statistic but does not reduce dispersion

about the mean, measured by the "residual mean squared". This result is an indication that the relationship may not be linear. Additional data and research is required to confirm that the relationships are non-linear and to develop a model suitable for long range requirements forecasting.

A TIME SERIES ANALYSIS OF RECOVERABLE SPARES REQUIREMENTS

I. INTRODUCTION

General Issue

Requirements for aircraft spare parts comprise a large portion of the Air Force budget. The requirement for peacetime operating stocks of recoverable replenishment spares (only a portion of the total aircraft spare parts requirement) totaled almost 7 billion dollars for fiscal years 1983, 1984, and 1985 according to a January 1983 AF/ACM study (7:A1). Accuracy in deriving forecasts for Budgets and Program Objective Memorandums (POMs) is paramount if the Air Force is to maintain credibility with Department of Defense, Office of Management and Budget, Congress, and others in the federal budgetary process. Historically our POM forecasts, projecting 3-7 years into the future, have been significantly understated. These understatements became increasingly apparent in late 1982 with the discovery of an understatement for fiscal years 1982-1984 totalling almost $900 million (5:56). The Air Force Chief of Staff subsequently ordered a comprehensive review of the spares forecasting process. This inquiry, called "Corona Require," identified numerous problems which combined to cause the understatements. Requirement computation system and methodology inadequacies were named as part of the problem (5:58).

Background

Inadequacies in the requirements computation system may be more than the cited "antiquated" computer being used (5:58). The current computation method assumes a linear causal relationship between flying hours for time t and demands for spares during time t. Forecasted demands become the basis for the spares budget requirement extending 2 years into the future. Prior to 1982, a budgeted cost per flying hour factor was multiplied by the forecasted flying hours to arrive at the more long term (3-7 years) POM requirement. However, in January 1983 AF/ACM sought to change the computation process and developed a technique to compute POM requirements at the macro level (dollar totals for a weapon system without visibility of item by item requirements). Linear regression techniques based on the age, value, and utilization of the aircraft fleet were used to develop a fiscal year 1983, 1984, and 1985 requirement for each major weapon system (i.e., C-141, F-16, C-130). This macro computation technique revealed POM understatements for fiscal years 1983, 1984, and 1985 of more than $1.3 billion (7:A1). This method, called Peacetime Operating Stock Spares Estimating System (POSSEM) (4), has been used as the basis for the POM submission since January 1983. Thus, the use of a linear regression which was previously questioned in item by item computations, remains a part of the new macro technique.

Specific Problem

A current expenditure for recoverable spares cannot support the current flying hour program. Rather, a current expenditure for recoverable spares supports a flying hour program after the spares are delivered (production leadtime away). This time lag between the expenditure and the support provided can be more than 3 years for some recoverable items. The current budgetary computation method ignores these time lags between the two variables. Hence, a budget requirement methodology which recognizes this lag may be more accurate than the current method.

Scope

This research develops a macro aircraft spares requirement forecasting technique using a Box - Jenkins time series analysis methodology. Here we discern relationships between aircraft expenditures (obligation) and flying hours. The terms obligation and expenditure are often used interchangeably. However, the term "obligation" is defined as a legally binding agreement to expend government funds for some future good or service. Since the Air Force budgets for obligation authority the more definitive term "obligations" will be used rather than "expenditures." In addition, the

research is limited to requirements and obligations for aircraft recoverable, replenishment spares (Budget Program 15000). Since the research is intended to demonstrate a methodology, this study will be limited to only one weapon system, the C-130 aircraft. The data base will include monthly obligation rates for C-130 unique items from January 1980 until March 1984 and C-130 monthly flying hours from January 1979 until March 1984.

Research Objectives

This research will investigate how Box - Jenkins time series analysis techniques can be used to develop a formula to forecast POM requirements for aircraft spares. This formula will then be compared with the POSSEM linear regression technique. This research is not intended to develop "the model" for long-range spares requirements forecasts, but rather to develop a greater understanding of the relationships between C-130 flying hours and C-130 spares requirements. Additional research will be required to examine relationships between spares requirements and other variables such as aircraft age, aircraft value, etc. Although a model will be developed based on relationships between flying hours and expenditures, a more accurate model will likely emerge only after study and inclusion of these additional relationships.

Availability of Data

Monthly obligations by weapon system were available from AFLC/ACB for the time period January 1980 thru March 1984. Cumulative monthly obligations for C-130 recoverable replenishment spares were obtained from the Financial Status of RDT&E and Procurement (H057.7J2R) reports. These reports are official accounting and finance reports and include all Air Force obligations for replenishment spares reported to the Department of Defense. The cumulative obligation data was converted to non-cumulative monthly data for further analysis resulting in 51 monthly data points. Monthly flying hours were obtained from the Air Force Inspection and Safety Center (AFISC/SERD). Non-cumulative monthly flying hours were collected for January 1979 through March 1984 for a total of 63 data points.

II. BACKGROUND

Types of Spares

Several classifications of spares are discussed in this paper. Spares can be either recoverable or consumable. In addition, spares are classified as either replenishment or initial spares. The research is limited to a review of recoverable, replenishment spares requirements. A recoverable spare is "a spare part which normally is not expended in use and which can be reused after recovery and repair (6:571)." Replenishment spares are spare "parts required for support of end items subsequent to the procurement of initial requirements (6:583)." Spare can also be classified as consumable and initial. A consumable spare or consumable item is "an item that is normally expended or used up beyond recovery in the use for which it was designed or intended (6:158)." Finally, an initial spare is "an item procured for logistics support of a system during its initial period of operation (6:348). Spares requirements are either recoverable or consumable and either initial or replenishment requirements.

Budget Process

The Budget Estimate Submission (BES) is submitted to DOD in October of each year for the fiscal year that begins the following October. For example, in October 1983 the Air Force submitted the BES for fiscal year 1985 which begins on 1 October 1984. The March recoverable consumption item requirements (D041) system computation (item by item) becomes the input for the Air Force BES submitted in October of each year. The BES submitted in October 1983 included not only requirements for fiscal year 1985 but also an update of the fiscal year 1984 requirement (4). The BES becomes the basis for the Program Objective Memorandum (POM) requirement which extends another 5 years into the future.

Program Objective Memorandum (POM)

The Air Force POM is submitted each year in January. The POM presents requirements for future years 3 through 7. For example, the fiscal year 1986 POM submitted in January 1984 included requirements for fiscal years 1986 through 1990. Prior to the January 1983 submission, the POM was primarily an extension of the BES adjusted for expected inflation and for changes in flying hour activity (4). The "Corona Require" inquiry stated that this practice could lead to an understatement of requirements.

> Say the program calls for a system to fly 1,000 hours, which will cause the consumption of ten widgets. There are nine widgets in stock. The projected cost per flying hour is the price of the one additional widget required, divided by 1,000.
>
> Then say there is a ten percent increase in flying hours. Total widget consumption goes up from ten to eleven, and the Air Force must buy two more widgets instead of one. The requirement has increased by 100 percent, but planners fine-tuning the budget may not realize this. They add ten percent more money to cover the ten percent increase in flying hours, and a deficit is created [5:58].

In an attempt to overcome computation inadequacies the January 1983 and 1984 POM submissions were developed using the macro techniques developed by AF/ACM in their 1982 study (7).

AF/ACM Study (POSSEM)

In October 1982, AF/ACM initiated a study in concert with AF/LEX and AF/PRP to

> analyze historical replenishment spares requirements and determine if an estimating technique could be developed at the macro level to forecast spares requirements [7:1].

The AF/ACM study used regression analysis to identify linear relationships between spares requirements for each weapon system (i.e., F-15, C-130, C-141) and the following variables.

1. Value of the fleet (including modification costs).

2. Utilization of the fleet.

3. Age of the fleet (actually 1/age was used).

The resulting equation for C-130 peculiar spares was "Requirement = .01184 (Value) - 853.2 (1/age) (C:A3/C-130)." C-130 utilization was not considered in the requirements equation since the study assumes that cargo aircraft spares requirements would not be as sensitive to utilization rates as fighters or bombers which have much more avionics equipment (C:A3/C-5)." This macro technique is called the Peacetime Operating Stock Spares Estimating Method (POSSEM) (4).

Air Logistics Early Requirements Technique (ALERT)

Mr. Jim Brannock, AFLC/MMMAA has proposed an alternative to POSSEM for forecasting long range POM requirements. This new procedure, called ALERT is also a linear regression model that computes POM requirements on a macro (by weapon system) basis (4). The primary difference between ALERT and POSSEM is the departure point used in developing the long-range forecast. POSSEM uses actual historical data. ALERT extends this actual data more than 2 years into the future and uses this micro-computed short-range forecast as if it were actual

data. Initial tests indicate that overall results of the two methods are not significantly different. However, these tests do indicate that ALERT provides a significant improvement when evaluated by weapon system (4). ALERT is expected to be used for the January 1985 POM submission.

Time Series Analysis

To generate a forecast of future events, the forecaster normally analyzes historical data concerning the event. Hopefully this analysis reveals some pattern in the occurrence of the event which can be extended into the future. This technique assumes that the identified pattern will continue into the future. The historical data used to prepare a forecast is often a time series. "A time series is a chronological sequence of observations on a particular variable (2:6)." In our case, the sequence of observations is monthly (i.e., total flying hours for a month or obligations for a month) and the variables under examination are spares obligations and flying hours. The historical patterns can include trends, cycles, seasonal variations, and/or irregular fluctuations. "Trend refers to the upward or downward movement that characterizes a time series over a period of time (2:7)." Trend is, therefore, the long-term rise or decline in the time series. "Cycle refers to the recurring up and down movement around trend levels (2:7)." A cycle, measured

from peak to peak or from trough to trough, lasts for 2 to 10 years. "Seasonal variations are periodic patterns in a time series that complete themselves within the period of a calendar year and are then repeated on a yearly basis (2:8)." Increased retail sales in December of each year is an example of seasonal variation. "Irregular fluctuations are erratic movements in a time series that follow no recognizable or regular pattern (2:8)." "Unusual" events that cannot be forecast such as fires, war, strikes, etc, can cause irregular fluctuations.

Box-Jenkins Methodology

The Box-Jenkins methodology is appropriate when the values of the time series under study are statistically dependent on one another. Other approaches such as regression assume that the values of the series are linearly independent of each other.

> If the values of the time series being forecasted are statistically dependent upon or related to each other, then the Box-Jenkins methodology uses the dependency to produce forecasts that are likely to be more accurate than the forecasts produced by the regression or exponential smoothing approaches [2:334].

The Box-Jenkins methodology provides a systematic procedure for analyzing and forecasting time series. Although the details of the method will be explained in the methodology

section of this paper, a brief overview provided by Bowerman and O'Connell follows:

> The Box-Jenkins methodology first develops an appropriate time series model for use in forecasting. This development consists of a three step iterative procedure. The first step is called identification. In this step a tentative model is identified by analysis of the historical data. The second step is called estimation. In this step the unknown parameters of the tentative model are estimated. The third step is called diagnostic checking. In this step diagnostic checks are performed to test the adequacy of the model, and if need be, to suggest potential improvements. Once a time series model has been developed, a fourth step, called forecasting, generates predictions of future values of the time series [2:335].

III. METHODOLOGY

Overview

This chapter will provide the methodology used to identify relationships within and between C-130 flying hours and obligations. First, data collection procedures for each time series will be explained. Then, each of the Box-Jenkins time series analysis steps (identification, estimation, and diagnostic checking) used to identify relationships within each time series will be discussed. Finally, the procedures to develop a transfer function and noise model to depict relationships between the two time series will be provided.

Data Collection - Obligations

Cumulative monthly obligations for C-130 spares were obtained from the Financial Status of RDT&E and Procurement (H057.7J2R) Report. This report is an official accounting and finance report which includes all Air Force obligations for replenishment spares reported to the Department of Defense. Monthly reports were available in AFLC/ACB dated from March 1984 back to 31 December 1979. Since reports dated prior to 31 December 1979 were retained in AFLC/ACB files only on a quarterly basis, monthly data prior to 31 December 1979 was not available. The cumulative obligation

data was converted to non-cumulative monthly data for further analysis. This data is attached as Appendix A.

Data Collection - Flying Hours

Monthly flying hours were obtained from the Air Force Inspection and Safety Center (AFISC/IGBM). Non-cumulative monthly flying hours for C-130 aircraft were collected for January 1979 through March 1984. Twelve additional months of flying hour data were collected as compared to the obligation data (which starts with January 1980) because of data transformations compatibility. This data is also attached as Appendix B.

Model Development

The Box-Jenkins time series analysis methodology suggests a step-by-step process to develop a model for forecasting a future value of the time series. These steps include (1) using historical data to identify a tentative model, (2) estimating the parameters of the tentative model, (3) improving the tentative model by accomplishing various diagnostic checks, and finally (4) using the resulting model to forecast future time series values. Since each of these steps is a complex and rigorous process, each step will be individually discussed.

The above procedures apply when analyzing both univariate (only one variable) and multivariate (more than one variable) time series. Since this review deals with a multivariate time series (both flying hours and obligations) relationships both within each time series and between the two series were identified. Transfer function (multivariate time series) modeling was accomplished to investigate relationships between the two time series. Transfer function modeling is the process of identifying and modeling the relationships between two time series.

> It is assumed that the dependent series, denoted here by y_t, can be represented by a linear operation on the independent series, x_t, plus a noise component n_t, where y_t, x_t, and n_t are all stationary. In general, x_t and y_t will be generated by differencing and possibly transforming the actual data series X_t and Y_t. Hopefully, additional information available from the X_t series can be used to obtain better forecasts of Y_t than would be the case were Y_t treated alone [12:15].

An example of a univariate model is $Y_t = .40Y_{t-1} + e_t$. An example of a multivariate model is $Y_t = 100 + .40X_{t-1} + N_t$. In this multivariate transfer model, the value of Y_t is dependent on the value of X in the prior time period, t-1.

Tentative Model Identification

Historical data for each time series (obligations and flying hours) were separately analyzed to identify a tentative model. First, a simple plot of each time series was prepared. A preliminary examination of these graphs provided an indication of the stationarity of each time series. A time series is stationary if its values fluctuate around a constant mean with a constant variance. If a time series is nonstationary, its values do not fluctuate around a constant mean. A nonstationary time series must be transformed to a stationary time series before identifying a tentative model. Stationarity can sometimes be induced by a transformation of the historical time series data known as first differencing.

If the time series observations are Y_1, $Y_2, \ldots. Y_n$, then the time series of the first differences will be X_1, $X_2, \ldots\ldots,$ X_{n-1} where $X_i = Y_i - Y_{i-1}$ for $i=2,\ldots.n$. If this process does not provide a stationary time series, taking the second differences ($W_i = Y_i - 2Y_{i-1} + Y_{i-2}$ for $i=3,\ldots.n$) of the original observations will usually provide a stationary time series. However, an analysis of both C-130 obligations and C-130 flying hours data indicate that both time series appear to be stationary. Therefore, differencing was not required.

An important characteristic of a stationary time series is that the statistical properties of the time series are unaffected by a shift of the time origin (2:341). In other words, if our flying hour data is stationary the relationships identified and our level of confidence in these relationships should remain constant whether early data or later data is analyzed.

Autocorrelation is an important statistical relationship that is used in part to test the stationarity of the time series data, to describe the pattern of the time series data, and to identify a tentative model for the time series. An autocorrelation coefficient at lag k is a measure of the linear relationship between any two time series observations separated by k units of time (a lag). An autocorrelation coefficient value is always between 1 and -1. A value near 0 means the two observations are linearly independent (not correlated) and a value near 1 or -1 means the two observations are linearly dependent and move together in a linear fashion (1, with a positive slope and -1, with a negative slope). For example, an autocorrelation coefficient near one at lag one means that the value of X_t is expected to be large (greater than the mean) if the value of X_{t-1} was also large. Conversely, an autocorrelation coefficient near -1 means the value of X_t is expected to be less than the mean

if X_{t-1} was large (more than the mean). An autocorrelation function (ACF), then is a plot of the autocorrelation coefficients at lag k for k = 1, 2,..., n-1. An ACF "dies down" if the values began to approach zero after a small number of lags. If, for example, the ACF at lag 1 equals .70. lag 2 equals .50, lag 3 equal .30, and then gets increasingly smaller at greater lags, the ACF is considered to "die down". If the autocorrelation cuts off or dies down fairly quickly, the time series is assumed to be stationary. Conversely, if the autocorrelation does not cutoff or dies down extremely slowly, the time series is assumed to be nonstationary. In this case, first differencing, second differencing, or some other transformation is required to generate a stationary time series.

Once the original time series data have been found to be stationary or have been transformed into a stationary time series, a tentative model can be identified. The tentative model will be identified by looking for typical patterns in the ACF and partial autocorrelation function (PACF). The PACF is similar to the ACF since both indicate the relationships between observations of the time series at different points in time (lags). The ACF at a particular lag can be influenced by a strong ACF at an adjacent lag. However, the PACF is not effected by autocorrelation of adjacent lags.

The PACF can, therefore, measures the strength of individual relationships. Patterns in the ACF and PACF will suggest the form of an autoregressive moving average, or autoregressive integrated moving average (ARIMA) model.

An ARIMA model recognizes the level of differencing, and the order of the autoregressive and moving average processes. Using Box-Jenkins notation, θ_i's represent the moving average coefficients, ϕ_i's represent the autoregressive coefficients, and e_i's represent the error portions of the tentative ARIMA model. The order of the model is denoted with subscripts p, d, and q with the model called an ARIMA (p, d, q) model. The d subscript is the level of differencing with a zero meaning no differencing was used. The subscripts p and q represent the order of the autoregressive and moving average portions of the model respectively. The number of spikes in the pattern of the autocorrelation function and the partial autocorrelation function corresponds to the values assigned p and q respectively. The order of the ARIMA model (values of p, d and q) identifies a tentative model form which can be expressed in a Box-Jenkins shorthand as an ARIMA (p, d, q) model. For example, an ARMIA (1,0,1) model can be written as $X_t = \phi_1 X_{t-1} - \theta_1 e_{t-1} + e_t$.

An ARIMA model may also be written using a backward shift operator B (3:8) where $BX_t = X_{t-1}$ or $B^m X_t = X_{t-m}$. The ARIMA (1,0,1) model then becomes $(1-\phi_1 B)\ X_t = (1-\theta_1 B)e_t$ using a backward shift operator:

Parameter Estimation

After identifying the tentative model, values for the parameters ϕ_i and θ_i must be estimated from the data. These estimates will be derived using the TIMES software package on the Harris computer at AFIT. The computer program proceeds through a search process using various estimates of the parameter values until those values are found which minimize the sum of the squares of the errors. These values are then used in the tentative model.

For example, given a tentative ARIMA (1,0,1) model of $X_t = \phi_1 X_{t-1} - \theta_1 e_{t-1} + e_t$, if final parameter values of .40 and .65 are estimated for ϕ_1 and θ_1 respectively, the model becomes $X_t = .40X_{t-1} - .65e_{t-1} + e_t$.

Diagnostic Checks

After the tentative model is identified and its parameters estimated, diagnostic checks are accomplished to test

the adequacy of the model. In general, a model is considered adequate if the residual differences between the individual time series observations and the forecast using the tentative model are white noise. The term white noise simply means that the residuals display no linear relationships between themselves and therefore cannot be forecast by any improvement of the model. The following eight diagnostic checks were completed to determine if the model was adequate (10).

1. A plot of the residuals is visually inspected for changes in mean and/or variance. The mean should be near zero and the mean and variance values should not be changing over time.

2. The ACF and PACF of the residuals should be near zero with no significant spikes. Any statistically significant spike would indicate the need for additional autoregressive or moving average parameters.

3. A Portmanteau "lack of fit" test should be accomplished. If the Q value is less than a chi-square value given a specified confidence level with the same degrees of freedom, then we cannot reject the null hypothesis that the the residuals are white noise. A 95% confidence level was selected for this research. If the Q value is less than the chi-square value, the residuals can be expected to be white noise at least 95% of the time.

4. The parameter values should not change over time. The time series data can be split into two parts and analyzed with the same model form as two separate (but shorter) time series. The estimated parameters should not differ significantly between the two.

5. Periodicity in the autocorrelation function of the residuals should not exist. Cumulative periodogram, "P_i" values, which represent a Fourier transformation of the autocovariances at period n/i, are computed. These values help in the identification of periodic autocorrelation. The cumulative periodogram should be checked to insure that all "P_i" values are within a 95% confidence bounds and in a fairly straight line. If the residuals are not white noise, the "P_i" values will not be "within bounds".

6. A histogram of the residuals should be reviewed to determine if a normal frequency distribution exists. The distribution of the residuals does not have to be normal; however, if probabilistic arguments are to be used then it is often useful for the distribution to be normal.

7. Review the variance of the residuals. In comparing alternative models, the model providing the lowest residual mean square is considered to be a more satisfactory model.

8. A power spectrum will be reviewed to ensure that additional seasonality does not remain in the residuals. The power spectrum, also a Fourier transformation of the data at period 1/n, should be fairly flat. A jump at .2 or 1/5 for example would indicate seasonality at the recipical period, or, for this case, at period 5. The power spectrum is similar to the cumulative periodogram.

If in any of these diagnostic checks, the residuals do not appear to be random, the Box-Jenkins process is repeated and a new tentative model is fitted to the data. This process continues until the final model produces residuals that are white noise. At this point the residual time series is considered to be prewhitened. This process was completed for both the flying hours and obligations times series.

Transfer Function Model

Up to this point, each of the univariate time series (flying hours and obligations) are separately analyzed to identify relationships within each time series and two prewhitened univariate time series models are developed. The residuals from the two prewhitened time series models are then analyzed to identify the relationships between the two time series. This analysis results in the development of a

transfer function model. A transfer function is a combination of the prewhitened independent and dependent univariant time series models (flying hours and obligations) plus the noise model. For example, a transfer function model might take the following general form.

$$(1 - \delta_1 B - \dots - \delta_r B^r)\, a_t = (w_0 - w_1 B - \dots - w_s B^s)\, b_{t-b} + N_t \qquad (1)$$

Where A_t is the prewhitened dependent series (called the output series) and b_t is the prewhitened independent series (known as the input series).

The noise and the transfer function models will be simultaneously determined using the same steps of identification, estimation, and diagnostic checking previously described. The noise model will deal with residuals from the prewhitened flying hours and obligations transfer function time series. The transfer function model deals with the relationship that the input series has on the output series in some prior or later time period. For example, a flying hour observation in one month may be effecting the obligations observed in some later or prior month. After adequate noise and transfer function models are determined, a model to forecast an obligation time series, called a final transfer model, is

developed by adding the noise model to the transfer function model.

Summary

The Box-Jenkins steps of identification, estimation, and diagnostic checking will be used to develop a number of models that explore relationships first within, and then between, the flying hours and obligations time series. These relationships will be combined into a final transfer function model to be used to forecast future obligation time series given an anticipated future flying hour time series. In developing the forecasting model the following questions will be answered.

-Are past C-130 spares obligations in a particular time period related to spares obligations in some prior time period?

-Are past C-130 flying hours in a particular time period related to spares obligations in some prior time period?

-Are there linear relationship between C-130 spares obligations and flying hours?

-Can these various relationships be used to develop an obligations time series forecasting model?

IV. ANALYSIS RESULTS

Overview

This chapter will provide the results of the Box-Jenkins time series analysis of monthly C-130 obligations and flying hours. Each step (identifying, estimating, and diagnostic checking) of the analysis of the two univariate time series will be discussed. Then, the results of the multivariate transfer function modeling will be addressed. In order to compare the Box-Jenkins forecasting techniques with the simple linear regression methods currently in use, a (0,0,0) transfer function model was prepared using the TIMES code. The results of this comparison also follows. In summary, the analysis revealed interrelationships within and between the two time series. Including these interrelationships in the forecasting model should provide more accurate forecasts. However, due to the scant data and the assumption of linear relationships, this analysis does not conclusively demonstrate Box-Jenkins techniques to be more accurate than simple linear regression techniques.

Obligations-Identification

Obligation data for 51 months (Appendix A) were analyzed using the TIMES computer code. First, a plot of the deviation from the mean (figure 1) was reviewed to obtain indications of stationary. Since the mean and the variance seem to be unchanging over time, the series appears stationary. The plot, however, reveals a very erratic obligation rate from month to month with no recognizable patterns. This is further confirmed by the chi-square value of 13.4 with 26 degrees of freedom which suggests that the original monthly obligations times series may be white noise. Plots of the ACF and PACF (figures 2 and 3) for lags 0 through 26 reveals weak moving average and possibly autoregressive patterns at lag 3. This seasonality at lag 3 is also confirmed by the power spectrum jump at .33 (figure 4). Since both the ACF and PACF values dampen quickly, stationarity in the series is confirmed. Therefore, parameters for three possible ARIMA models of the form $(0,0,0)*(0,0,1)_3$; $(0,0,0)*(1,0,0)_3$; and $(0,0,0)*(1,0,1)_3$ were estimated. In summary, the identification process revealed no significant serial interrelationships within the obligation series. However, weak seasonality patterns at lag 3 were evident and were investigated further.

DEVIATIONS FROM MEAN OBLIGATIONS

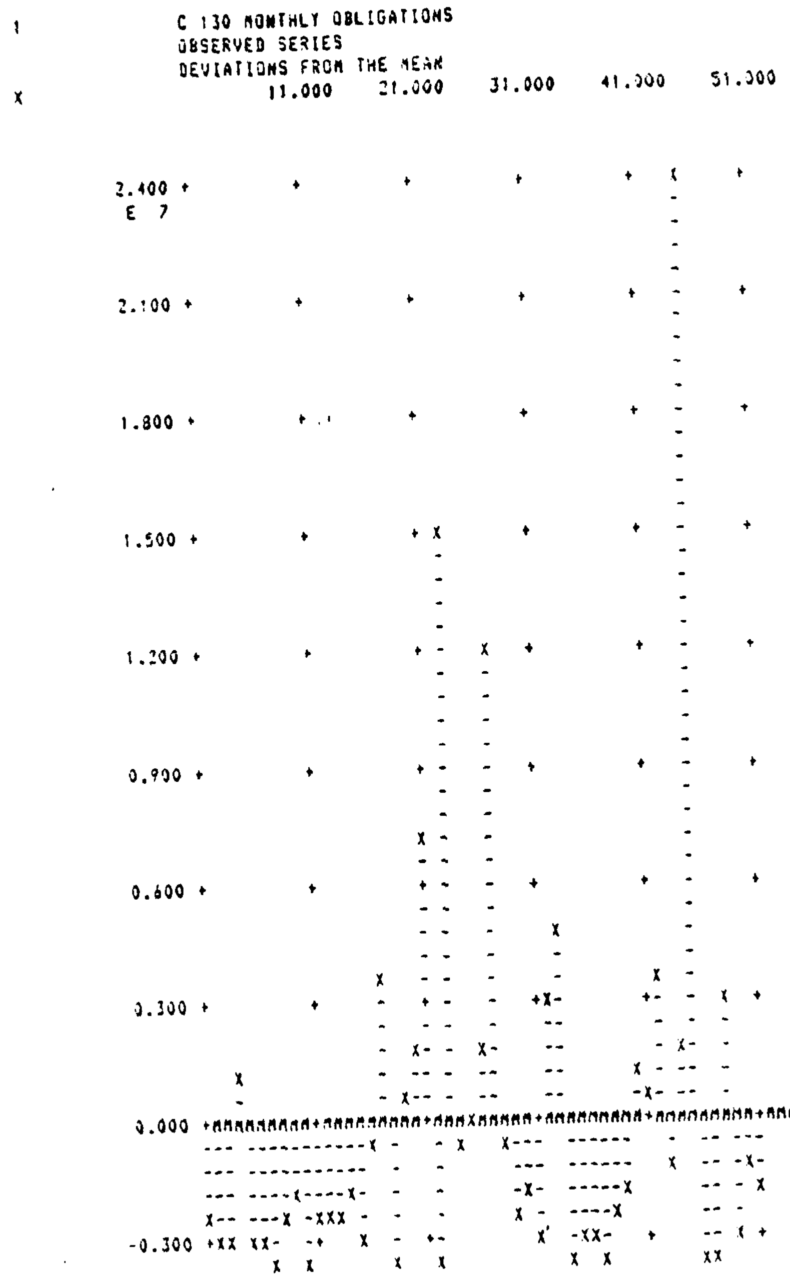


Figure 1. Deviations from Mean-Obligations

AUTOCORRELATION FUNCTION
OBLIGATIONS

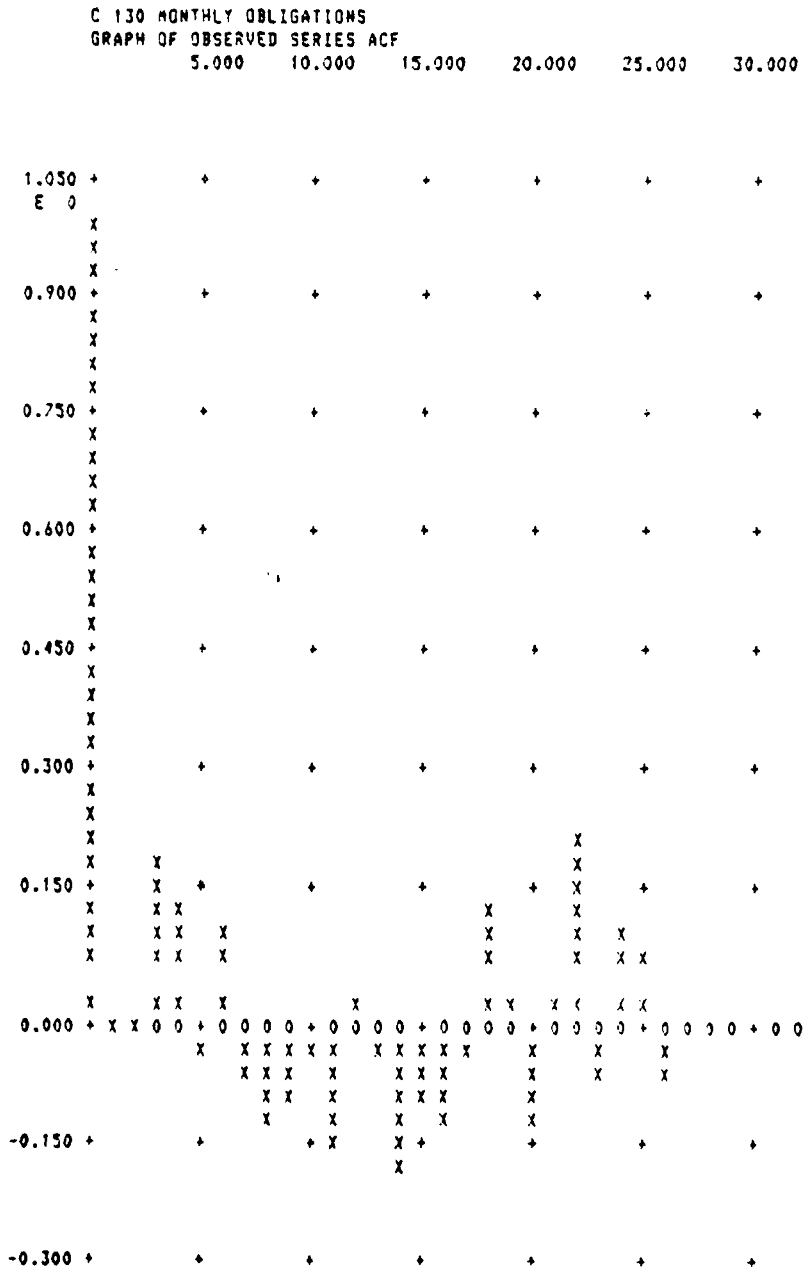


Figure 2. Autocorrelation Function - Obligations

PARTIAL AUTOCORRELATION FUNCTION
OBLIGATIONS

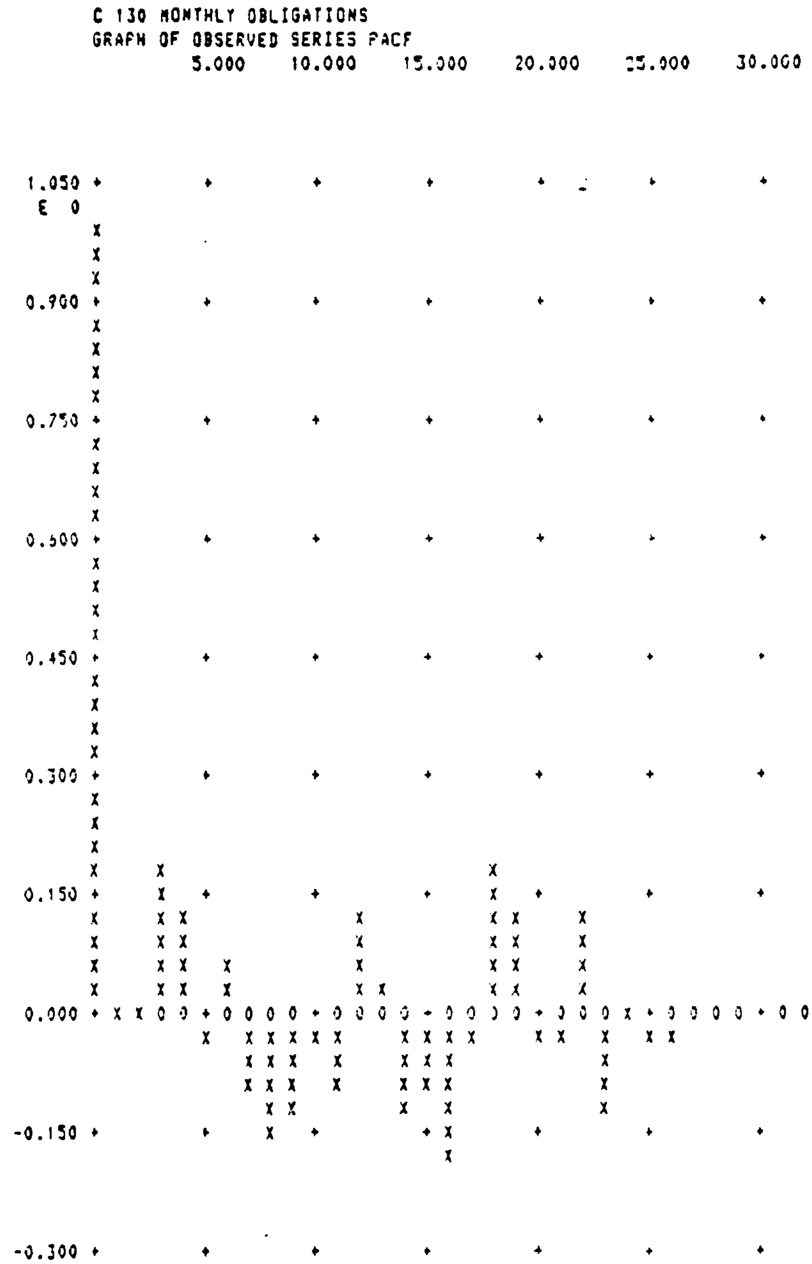


Figure 3. Partial Autocorrelation Function - Obligations

POWER SPECTRUM
OBLIGATIONS

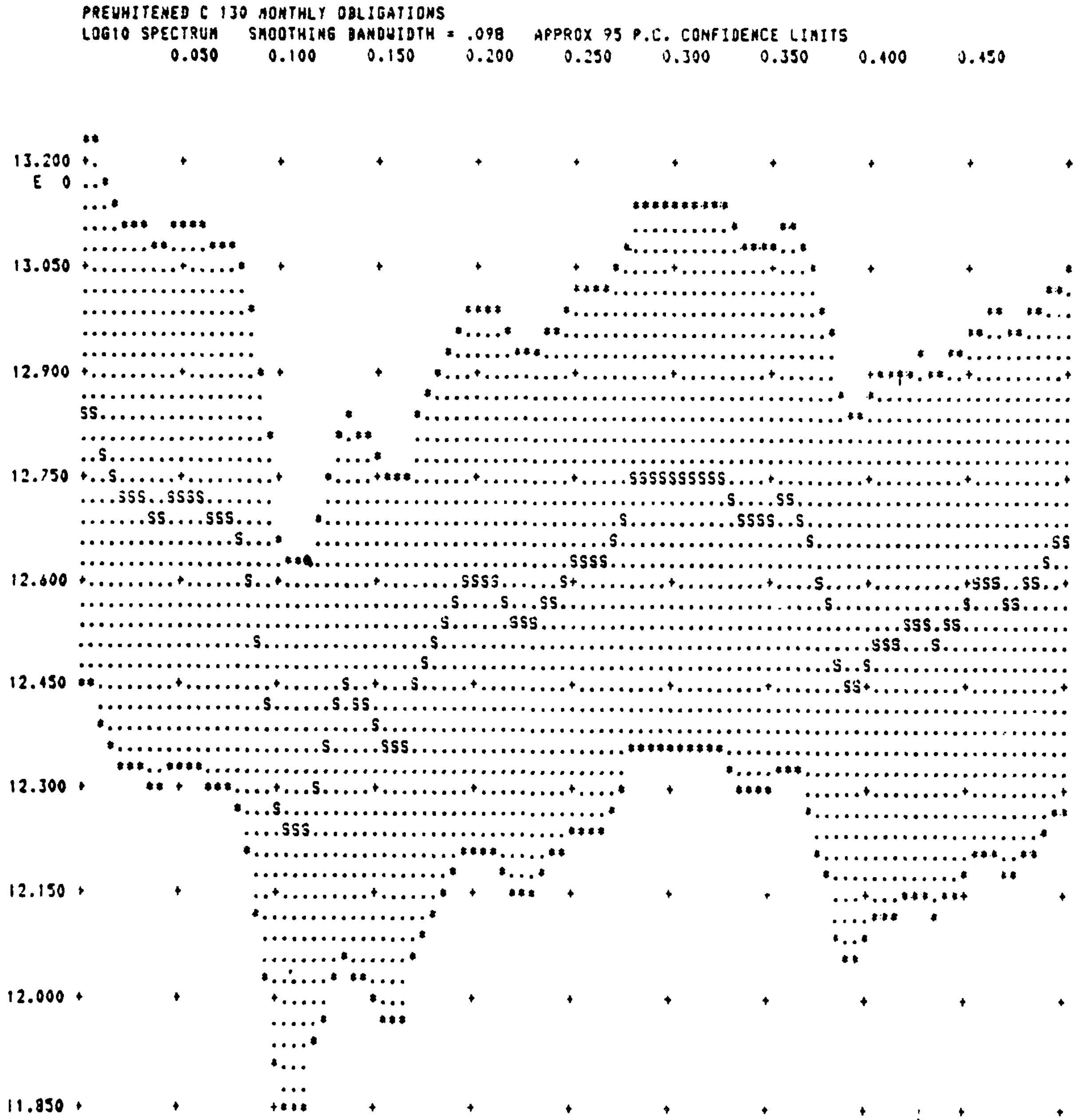


Figure 4. Power Spectrum - Obligations

ARIMA Model	Residual Mean Square
$(0,0,0)*(0,0,1)_3$	.26249 E+14
$(0,0,0)*(1,0,0)_3$	.27322 E+14
$(0,0,0)*(1,0,1)_3$	.27390 E+14

Figure 5. Residual Mean Square Comparisons - Obligations

<u>Obligations</u> - <u>Estimation</u>

Parameter values for ARIMA models $(0,0,0)*(0,0,1)_3$, $(0,0,0)*(1,0,0)_3$ and $(0,0,0)*(1,0,1)_3$ were estimated using the times package. The TIMES package estimates coefficient values which provide the lowest residual mean squares value. As reflected in figure 5, an ARMIA $(0,0,0)*(0,0,1)_3$ model provided the lowest residual mean square. This model also produced a lower chi-square value than the other alternatives. The small coefficient value - .14239 of the moving average parameter with a 95% confidence interval ranging from - .42383 to .13905 (see figure 6) confirmed that the interrelationship with the series is relatively weak. The chi-square statistic of 11.67 with 24 degrees of freedom obtained using the ARIMA $(0,0,0)*(0,0,1)_3$ model is an improvement over the chi-square value of 13.40 with 26 degrees of freedom from the original series. Therefore, an ARMIA $(0,0,0)*(0,0,1)_3$ model was selected for further diagnostic checks.

SUMMARY OF THE MODEL
OBLIGATIONS

SUMMARY OF MODEL 1

****

DATA - Z = C 130 MONTHLY OBLIGATIONS — 51 OBSERVATIONS

DIFFERENCING ON Z - NONE

PARAMETER NUMBER	PARAMETER TYPE	PARAMETER ORDER	ESTIMATED VALUE	95 PER CENT LOWER LIMIT	95 PER CENT UPPER LIMIT
1	MEAN	0	0.36307E+07	0.19788E+07	0.52827E+07
2	MOVING AVERAGE 1	3	-.14239E+00	-.42383E+00	0.13905E+00

OTHER INFORMATION AND RESULTS

RESIDUAL SUM OF SQUARES	0.12862E+16	49 D.F.	RESIDUAL MEAN SQUARE	0.26249E+14
NUMBER OF RESIDUALS	51		RESIDUAL STANDARD ERROR	0.51234E+07

Figure 6. Summary of the Model - Obligations

Obligations - Diagnostics

A visual inspection of the plot of deviations from the residual mean (figure 7) reveals a mean which is near zero and unchanging over time. The variance from the means is also unchanging over time. Both are indications of a satisfactory model.

Next, plots of the ACF and PACF (figures 8 and 9) were examined for significant (greater than .28) spikes. Bartletts rule was used to compute the two standard error bounds of + or - .28. According to Bartletts rule, one standard error can be approximated using the formula (1/n) (10). Two standard errors (or a 95% confidence interval if the residuals are normal) was used to evaluate the ACF for significant (non-zero) spikes. A significant spike could indicate the need for an additional parameter in the model. The largest ACF spike valued at .21 present at lag 22 was not considered significant partly because it was less than two standard deviations and partly because the estimated value of two standard deviations becomes less reliable due to the low number of data points being evaluated at larger lags. The lack of significant spikes in the ACF and PACF indicates that the model cannot be significantly improved by introducing additional parameters.

DEVIATIONS FROM MEAN
OBLIGATIONS RESIDUALS

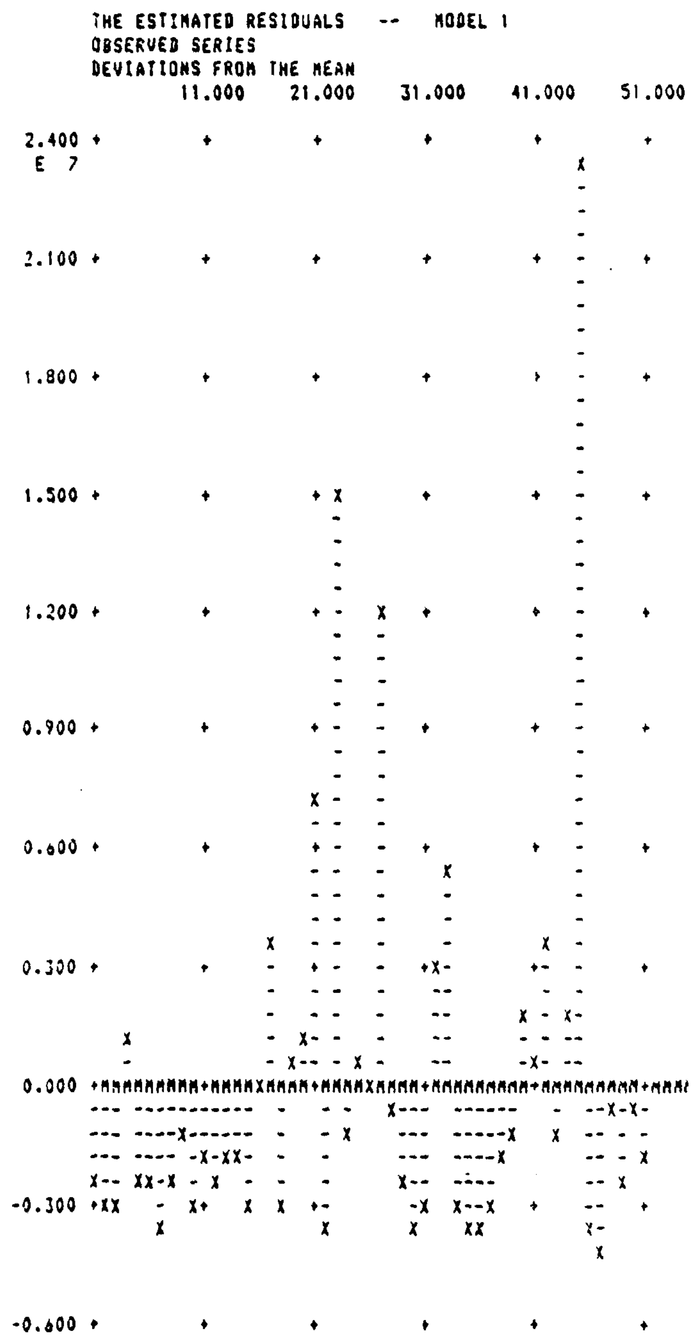


Figure 7. Deviations from Mean - Obligations Residuals

AUTOCORRELATION FUNCTION
OBLIGATION RESIDUALS

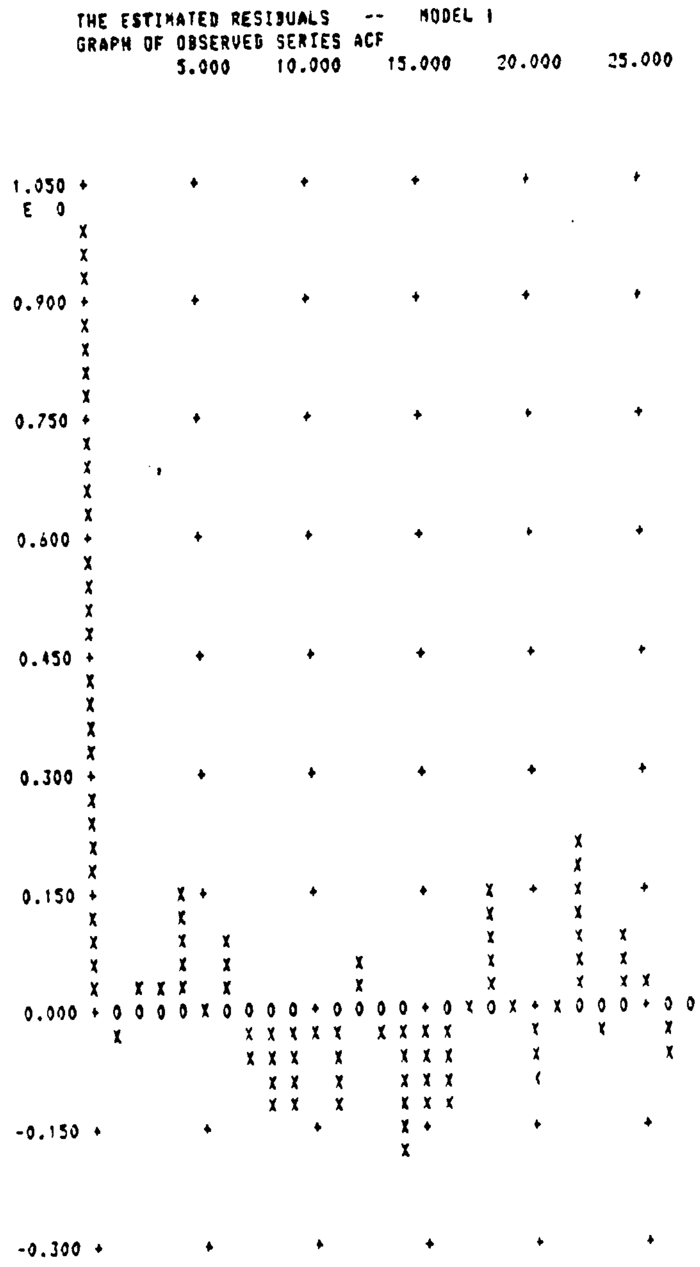


Figure 8. Autocorrelation Function - Obligation Residuals

PARTIAL AUTOCORRELATION FUNCTION
OBLIGATION RESIDUALS

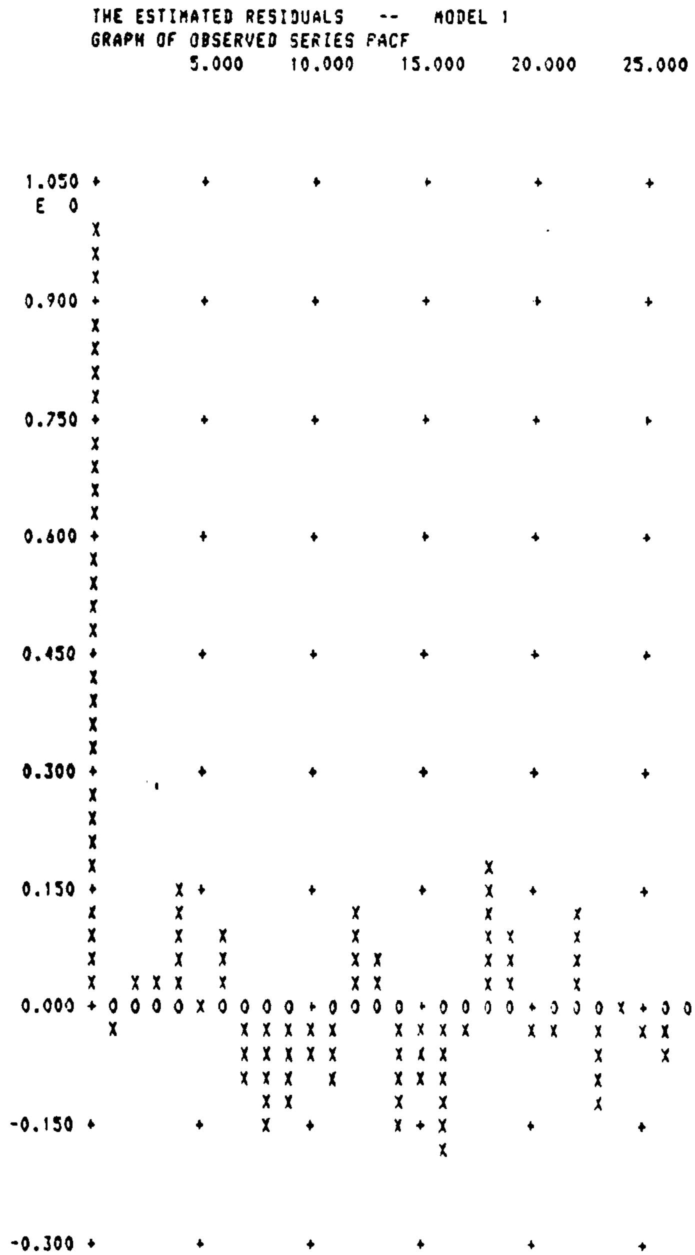


Figure 9. Partial Autocorrelation Function - Obligation Residuals

Next, a "Portmanteau "lack of fit" test (3:290) was accomplished. Since the Q statistic of 11.67 with 24 degrees of freedom is less than the $X^2_{.05}$ value of 36.415 (1:899), the residuals appear to be white noise, and hence, are not autocorrelated. The cumulative periodogram and power spectrum (figures 10 and 11) also confirm that the residuals are white noise. The histogram of the estimated residuals (figure 12) reveals a frequency distribution which is not normal. However, since the probability of a residual being a certain value is not required, a normal distribution is not a necessity. In summary, an ARIMA $(0,0,0)*(0,0,1)_3$ model with the moving average parameter valued at -.14239 and a mean of $3,630,700 provided the lowest residual mean square and provided residuals which were white noise. These residuals became the prewhitened obligation series (a_t) used in identifying the final transfer function and noise models.

Flying Hours - Identification

Originally, monthly flying hour data for 52 months were analyzed. This initial analysis, however, indicated strong seasonal autoregressive and/or moving average parameters at lag 12. The seasonality resulted in a reduction from the original 52 data points to 40 residual data point. Since

CUMMULATIVE PERIODOGRAMS
OBLIGATION RESIDUALS

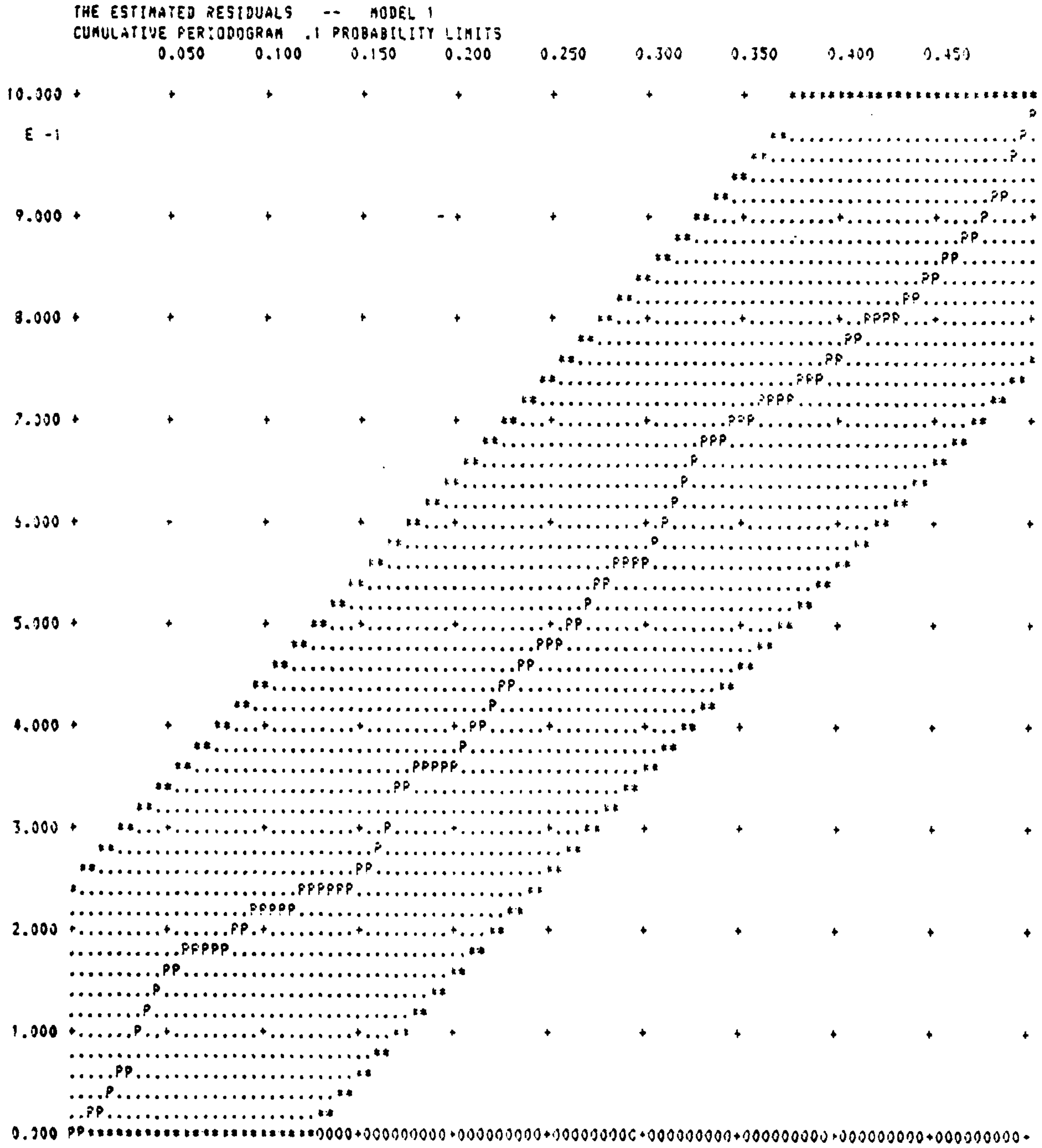


Figure 10. Cummulative Periodograms - Obligation Residuals

POWER SPECTRUM
OBLIGATION RESIDUALS

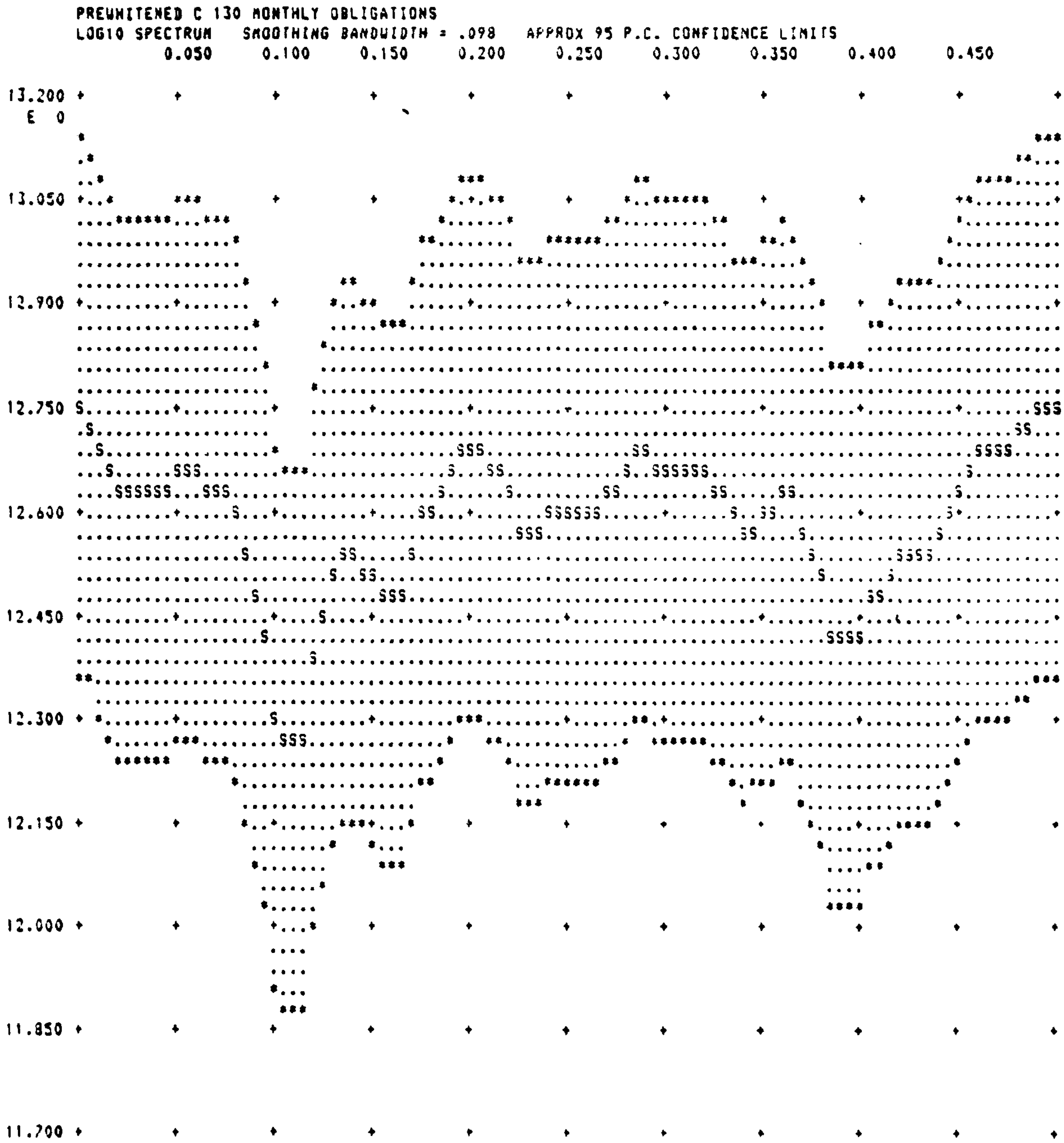


Figure 11. Power Spectrum - Obligation Residuals

HISTOGRAM
OBLIGATION RESIDUALS

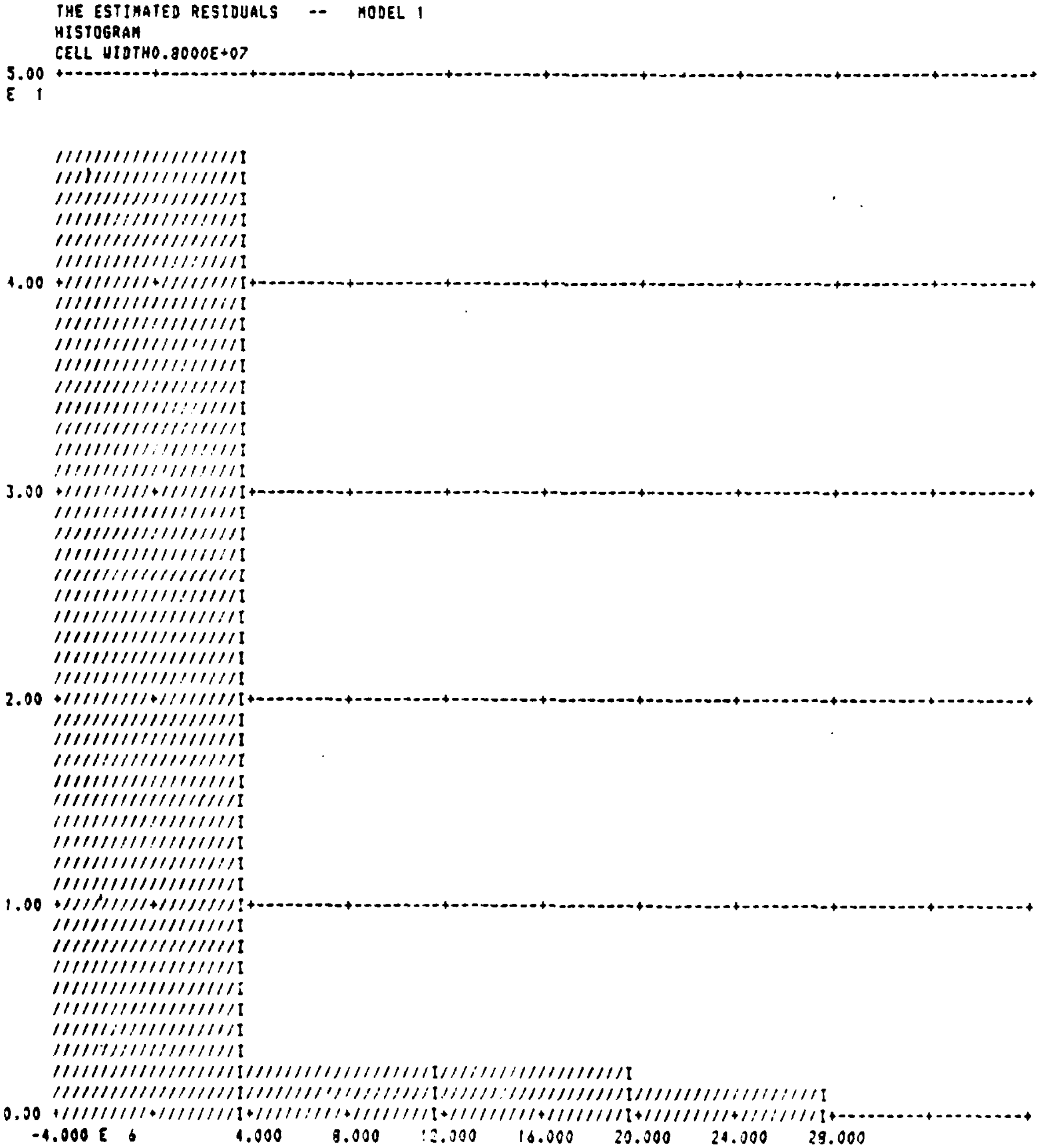


Figure 12. Histogram - Obligation Residuals

additional flying hour data were available, the number of months being analyzed was increased to 63 in order to facilitate a comparison with prewhitened obligations. The first step in the analysis was a determination of the stationarity of the flying hour data. A review of the data (Appendix B) and the plot of deviations from the mean (figure 13) indicates a stationary series since the mean and variance do not change over time. Both the ACF and PACF (figures 14 and 15) values dampen quickly (except for seasonality effects) and further confirm the stationarity of the data. The original data is not white noise since the chi-square value of 94.6 with 24 degrees of freedom is greater than a $X^2{.05}$ value of 36.415 (1:899). The ACF and PACF plots also confirm the existence of patterns in the data which indicate time dependencies. These patterns suggest a serial and seasonal ARIMA model of the form $(0,0,1)*(1,0,1)_{12}$. Parameter values for this model and for an extension of this model to an ARIMA $(0,0,1)*(0,0,1)_6*(0,0,1)_9*(1,0,0)_{12}$ model were estimated.

Flying Hours - Estimation with 52 Data Points

Using the original 52 data points (before expansion to 63 data points) parameters for an ARIMA $(0,0,1)*(1,0,1)^{12}$ model were first estimated. The resulting ACF and PACF

DEVIATION FROM THE MEAN
FLYING HOURS

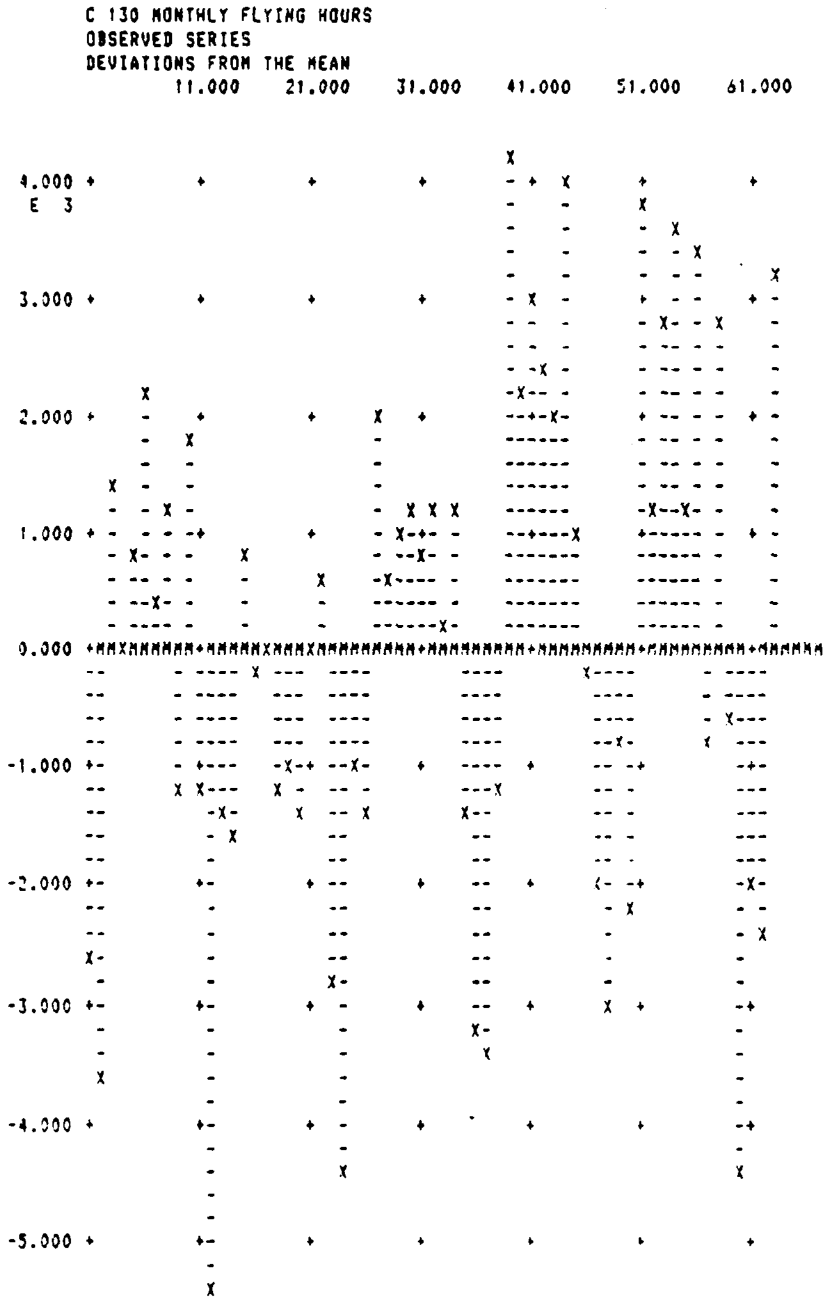


Figure 13. Deviation from the Mean - Flying Hours

AUTOCORRELATION FUNCTION
FLYING HOURS

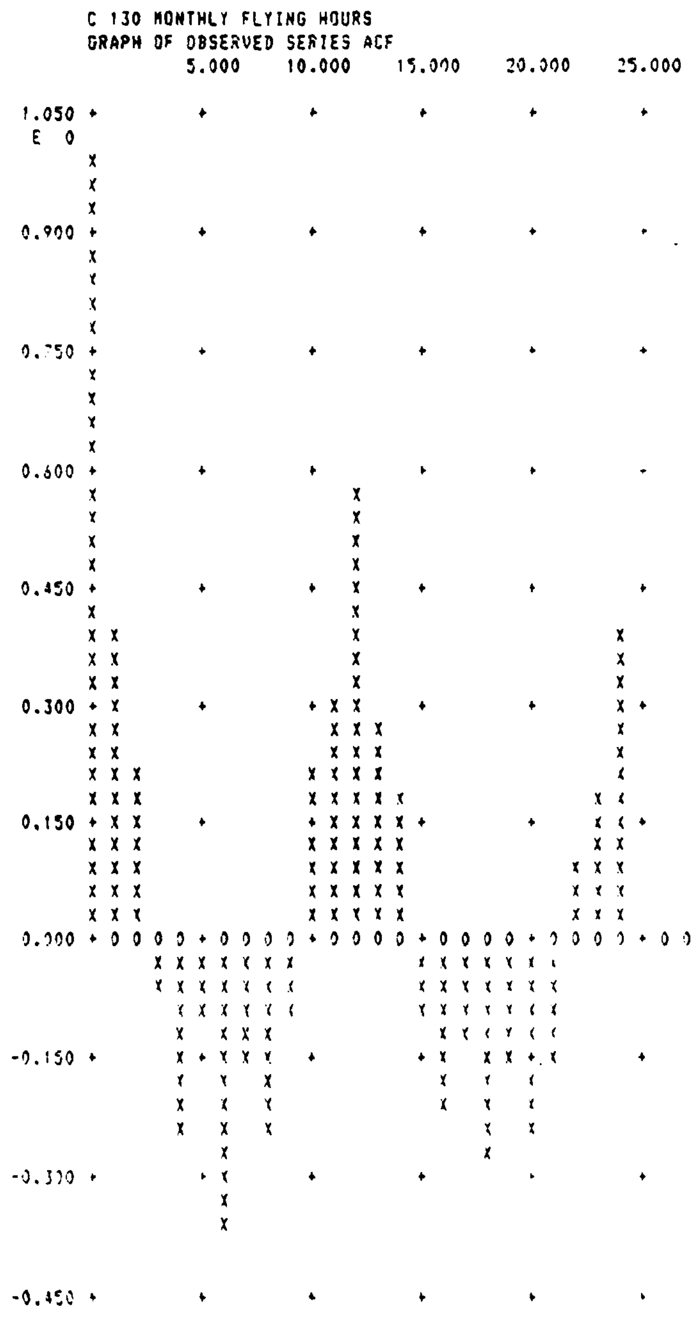


Figure 14. Autocorrelation Function - Flying Hours

PARTIAL AUTOCORRELATION FUNCTION
FLYING HOURS

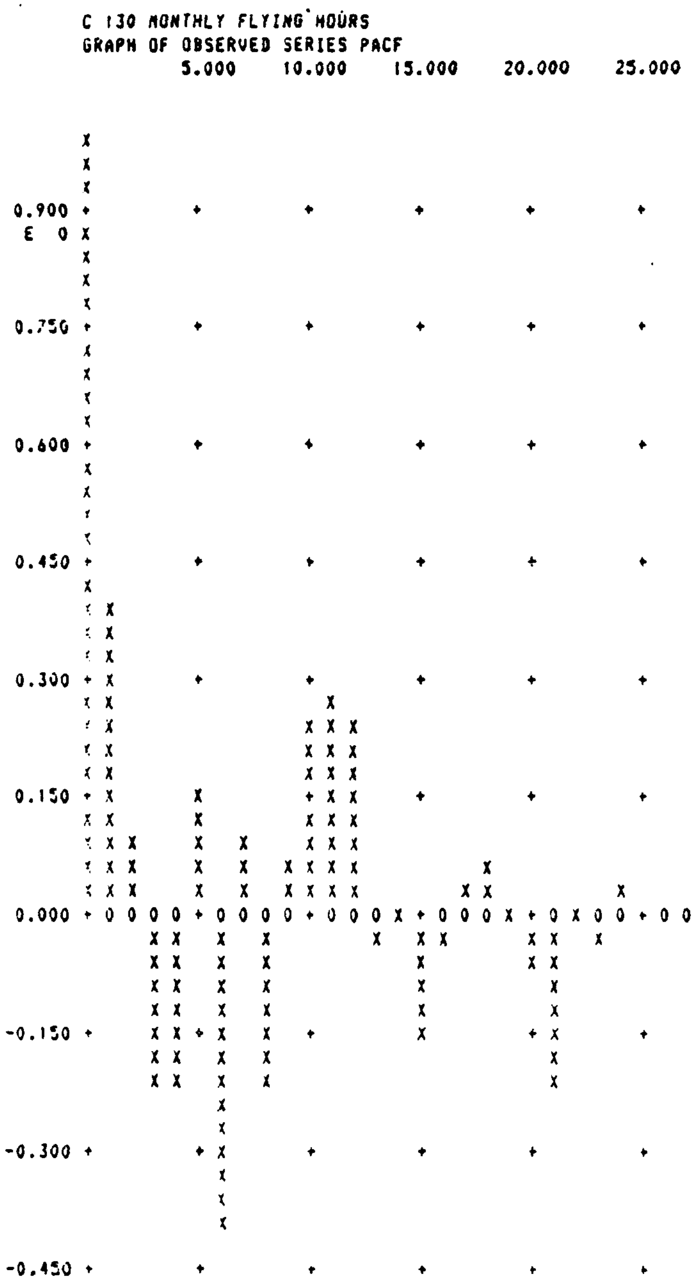


Figure 15. Partial Autocorrelation Function - Flying Hours

AUTOCORRELATION FUNCTION
FLYING HOUR RESIDUALS

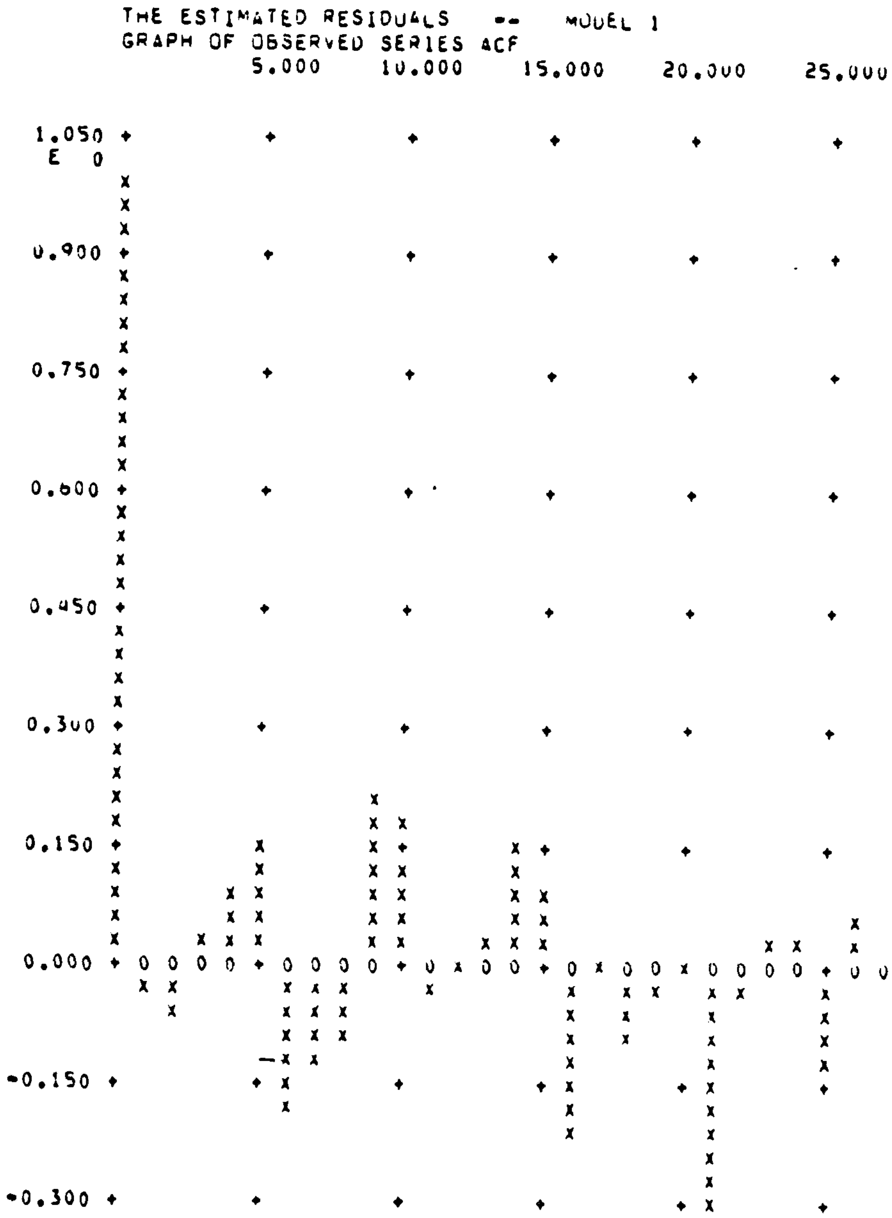


Figure 16. Autocorrelation Function - Flying Hour Residuals

ARIMA Model	Chi-Square	Residual Mean Squared
$(0,0,1)*(1,0,1)_{12}$	14.4 @ 22 d.f.	.172 E+7
$(0,0,1)*(0,0,1)_9(1,0,1)_{12}$	16.6 @ 21 d.f.	.149 E+7
$(0,0,1)*(0,0,1)_6*(1,0,1)_{12}$	12.1 @ 21 d.f.	.152 E+7
$(0,0,1)*(0,0,1)_6*(1,0,0)_{12}$	12.4 @ 22 d.f.	.150 E+7
$(0,0,1)*(0,0,1)_6*(0,0,1)_9*(1,0,0)_{12}$	10.4 @ 21 d.f.	.128 E+7
$(0,0,1)*(1,0,0)_3*(1,0,0)_{12}$	15.6 @ 22 d.f.	.187 E+7
$(0,0,1)*(1,0,0)_3$	62.6 @ 23 d.f.	.452 E+7
$(0,0,1)*(0,0,3)_3$	40.8 @ 21 d.f.	.469 E+7
$(0,0,1)*(0,0,2)_3*(1,0,0)_{12}$	13.1 @ 21 d.f.	.144 E+7
$(0,0,1)*(0,0,1)_3*(1,0,0)_{12}$	16.0 @ 22 d.f.	.172 E+7

Figure 18. Flying Hour Estimation Results with 52 Data Points

(figures 16 and 17) of the residuals indicated additional seasonal moving average parameters at lags 6 and 9. Because of the seasonal parameters at multiples of three, the values for the following four, more parsimonious, ARIMA models were also estimated:

$(0,0,1)*(1,0,0)_3*(1,0,0)_{12}$
$(0,0,1)*(1,0,0)_3$
$(0,0,1)*(0,0,3)_3$
$(0,0,1)*(0,0,1)_3*(1,0,0)_{12}$

The chi-square statistic and residual mean square comparison are provided in figure 18. Of the models examined using 52 data points, an ARIMA $(0,0,1)*(0,0,1)_6*(0,0,1)_9*(1,0,0)_{12}$ provided the best residuals mean squared and chi-square results. However, due to the low number of flying hour residuals (40) provided compared to the number of obligation residuals (51), the analysis data base was expended to include 63 months of flying hour data.

PARTIAL CORRELATION FUNCTION
FLYING HOUR RESIDUALS

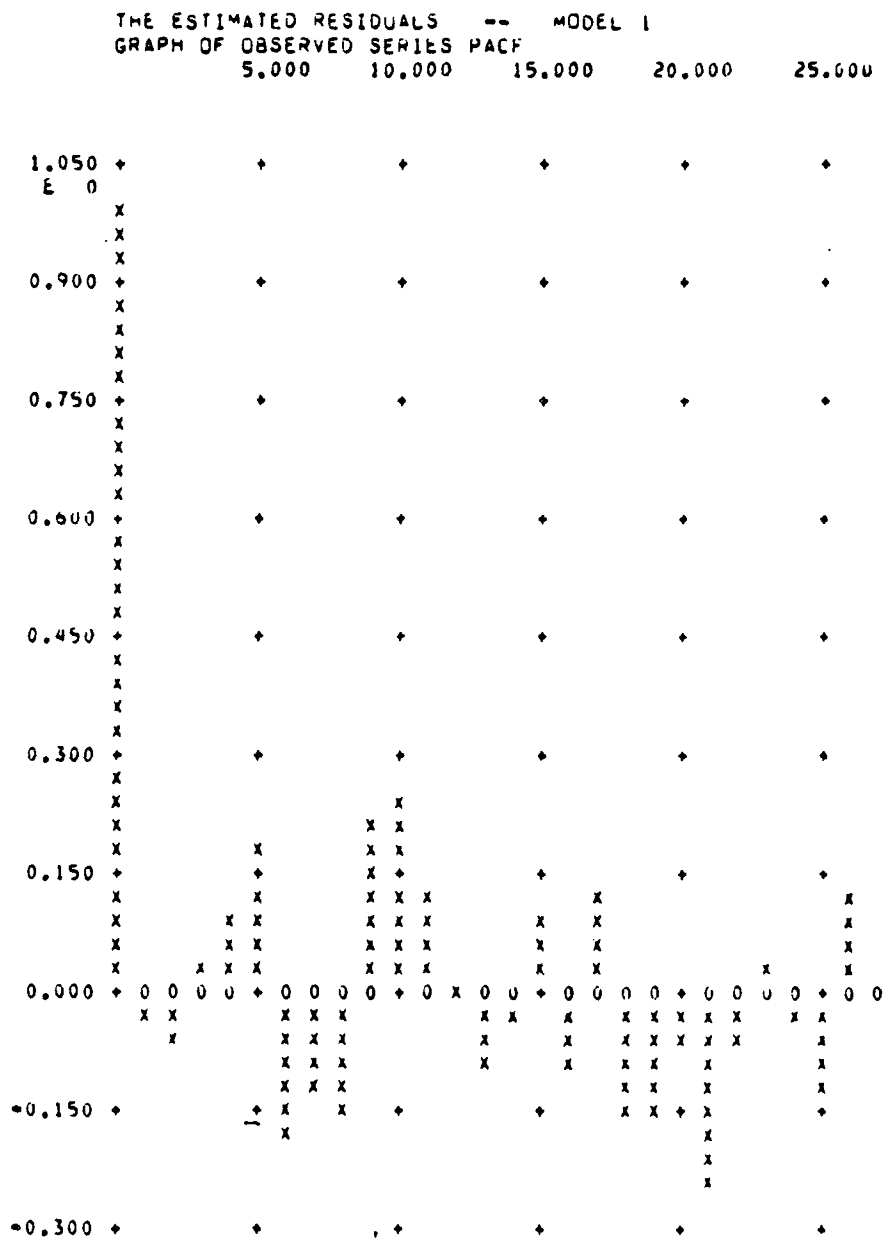


Figure 17. Partial Correlation Function - Flying Hour Residuals

ARIMA Model	Chi-Square	Residual Mean Squared
$(0,0,1)*(0,0,1)_6*(1,0,0)_{12}$	21.9 @ 22 d.f.	.177 E+7
$(0,0,1)*(0,0,1)_5*(0,0,1)_6*(1,0,0)_{12}$	16.3 @ 21 d.f.	.169 E+7
$(0,0,1)*(0,0,1)_6*(0,0,1)_9*(1,0,0)_{12}$	18.3 @ 21 d.f.	.175 E+7

Figure 19. Flying Hour Estimation Results with 63 Data Points

Flying Hours - Estimation with 63 Data Points

Parameter values for the following three ARIMA models were estimated:

$$(0,0,1)*(0,0,1)_6*(1,0,0)_{12}$$
$$(0,0,1)*(0,0,1)_5*(0,0,1)_6*(1,0,0)_{12}$$
$$(0,0,1)*(0,0,1)_6*(0,0,1)_9*(1,0,0)_{12}$$

The chi-square and residual mean square results are reflected in figure 19. After estimating the parameter values of an ARIMA $(0,0,1)*(0,0,1)_6*(0,0,1)_9*$ $(1,0,0)_{12}$ model and analyzing the resulting ACF and PACF of the residuals, additional seasonality at lag 5 and a decrease in the magnitude of the seasonality at lag 9 became apparent. The additional 12 data points included in the analysis were older data. Therefore, using more current data (the original 52 data points) the seasonality at lag 5 (for which there is no

intuitive rationale) disappears and the seasonality at lag 9 becomes stronger. Since there is no reason to expect this trend to change in the future, an ARIMA $(0,0,1)*(0,0,1)_6$ $(0,0,1)_9(1,0,0)_{12}$ was selected as the best model for forecasting. The estimated results are reflected in figure 20. The results of diagnostic checks of this model follow.

Flying Hours - Diagnostics

A visual inspection of the plot of deviation from the means of the residuals from a $(0,0,1)*(0,0,1)_6*(0,0,1)_9*$ $(1,0,0)_{12}$ model (figure 21) reveals a mean which is near zero and stationary. The residuals appear to be stationary since the variance is also unchanging over time. A review of plots of the ACF and PACF (figures 22 and 23) of flying hour residuals for lags 1 through 15 reveals no significant (greater than .28) spikes. The spikes at lags 16 and 21 were not considered due to the small number of data points available for analysis. The ACF and PACF patterns confirm that further significant forecasting improvements are not possible through inclusion of additional parameters in the model. Next, a Portmanteam "lack of fit" test was accomplished. The Q statistic of 18.313 with 21 degrees of freedom, is less than the chi-square ($X^2_{.05}$) value of 31.671 confirming that the residuals are white noise and not autocorrelated.

SUMMARY OF THE MODEL
FLYING HOURS

SUMMARY OF MODEL 1

DATA - Z = C 130 MONTHLY FLYING HOURS 63 OBSERVATIONS

DIFFERENCING ON Z - NONE

PARAMETER NUMBER	PARAMETER TYPE	PARAMETER ORDER	ESTIMATED VALUE	95 PER CENT LOWER LIMIT	95 PER CENT UPPER LIMIT
1	AUTOREGRESSIVE 1	12	0.79945E+00	0.62212E+00	0.97678E+00
2	MEAN	0	0.32109E+05	0.30034E+05	0.34185E+05
3	MOVING AVERAGE 1	1	-.37349E+00	-.65445E+00	-.92526E-01
4	MOVING AVERAGE 2	6	0.54114E+00	0.24670E+00	0.83558E+00
5	MOVING AVERAGE 3	9	-.22142E+00	-.53415E+00	0.91320E-01

OTHER INFORMATION AND RESULTS

RESIDUAL SUM OF SQUARES	0.80589E+08	46 D.F.	RESIDUAL MEAN SQUARE	0.17519E+07
NUMBER OF RESIDUALS	51		RESIDUAL STANDARD ERROR	0.13236E+04

Figure 20. Summary of the Model - Flying Hours

DEVIATIONS FROM THE MEAN
FLYING HOUR RESIDUALS

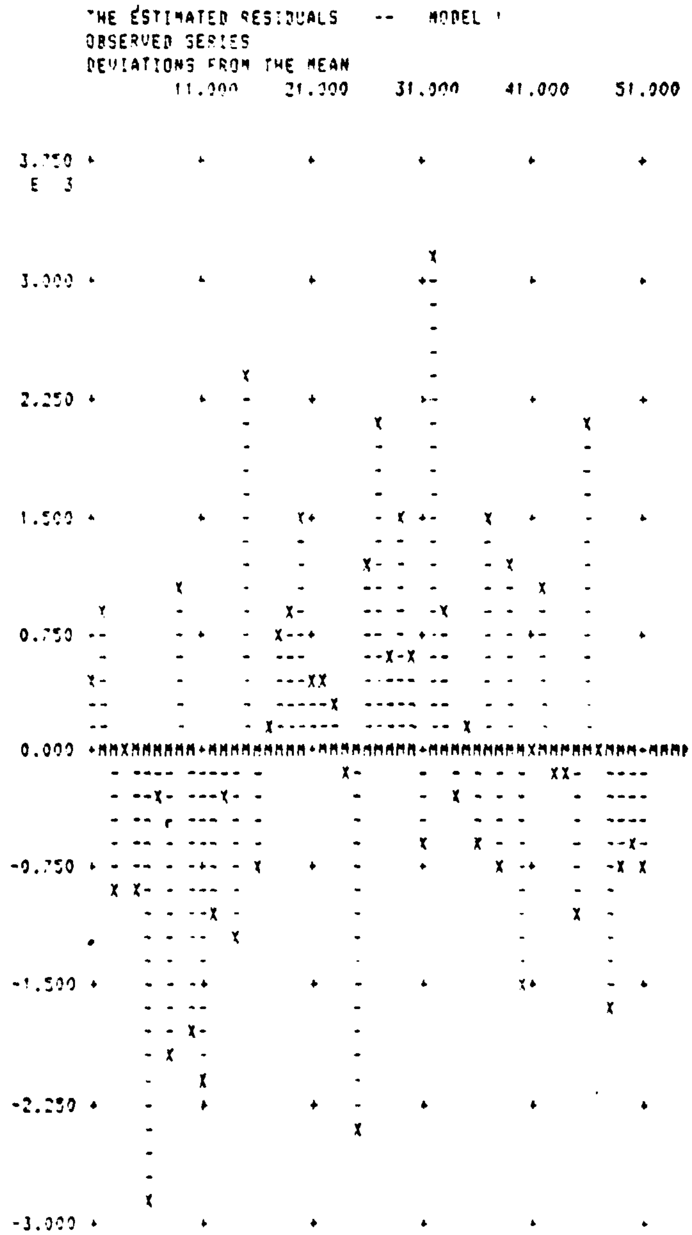


Figure 21. Deviations from the Mean - Flying Hour Residuals

**AUTOCORRELATION FUNCTION
FLYING HOUR RESIDUALS**

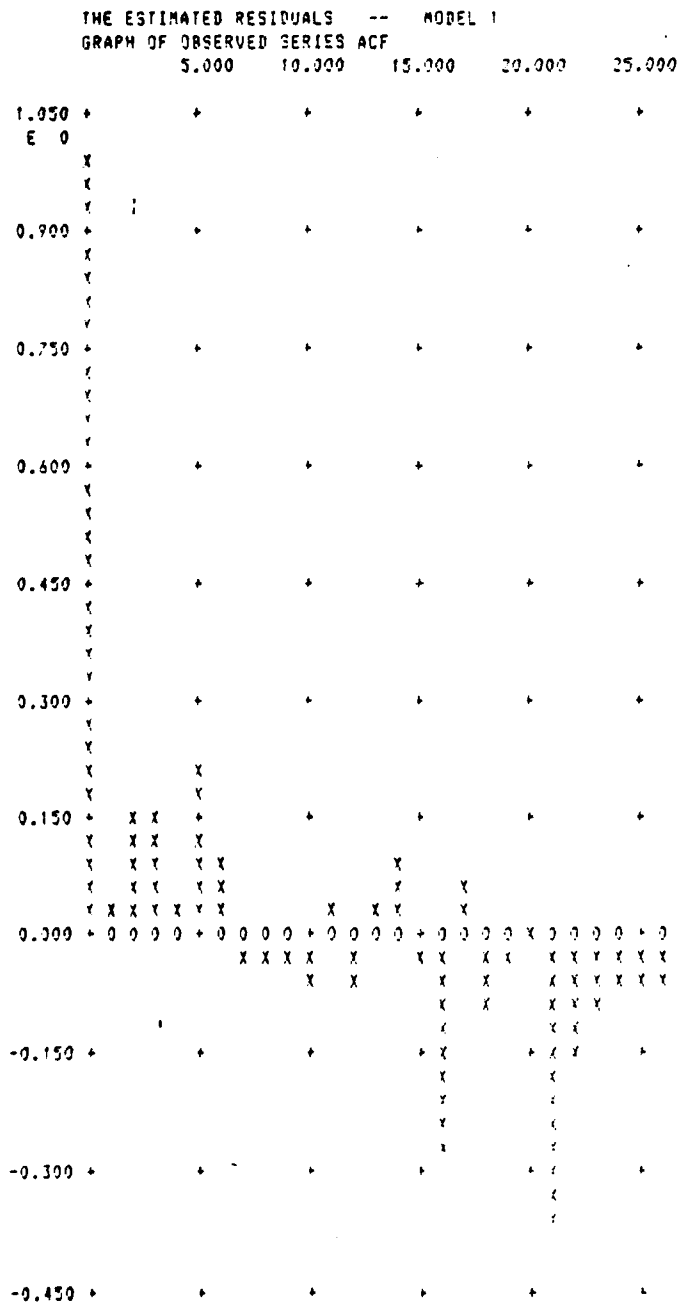


Figure 22. Autocorrelation Function - Flying Hour Residuals

PARTIAL AUTOCORRELATION FUNCTION
FLYING HOUR RESIDUALS

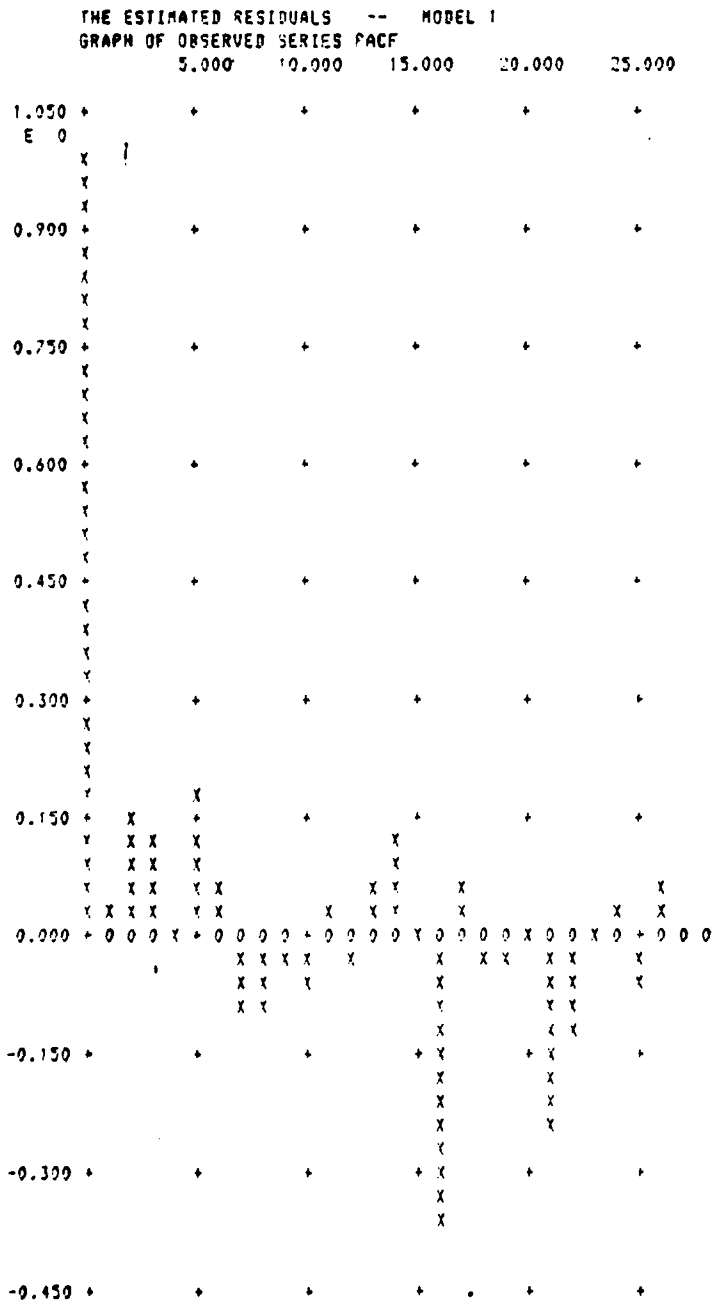


Figure 23. Partial Autocorrelation Function - Flying Hour Residuals

The cumulative periodogram and power spectrum (figures 24 and 25) also indicate that the residuals are white noise. The histogram of the estimated residuals (figure 26) reveals residuals with a mean near zero and a near normal frequency distribution. The flying hour data was reduced to the most current 39 data points to determine if the parameter values were changing over time. The new parameter values (figure 27) did not change significantly (within a 95% confidence interval) when using the last 39 data points. The ACF and PACF with the most recent 39 data points also confirmed that the magnitude of the seasonality at lag 5 is decreasing, while the parameter value for the seasonality at lag 9 also confirmed that the magnitude of lag 9 seasonality is increasing. In summary, an ARMIA $(0,0,1)*(0,0,1)_6*(0,0,1)_9*(1,0,0)_{12}$ model is a satisfactory univariate model for C-130 monthly flying hours. The residuals, b_t, from this model became the prewhitened flying hours series used in identifying the final transfer function and noise models.

Transfer Function/Noise Model Identification

The identification of possible transfer function forms was accomplished by analyzing values and patterns of the plots of cross correlations (figure 28). Significant cross correlation (values exceeding .28 in this example using two standard errors) occurs at lag 20. The cross correlation at

CUMMULATIVE PERIODOGRAM
FLYING HOUR RESIDUALS

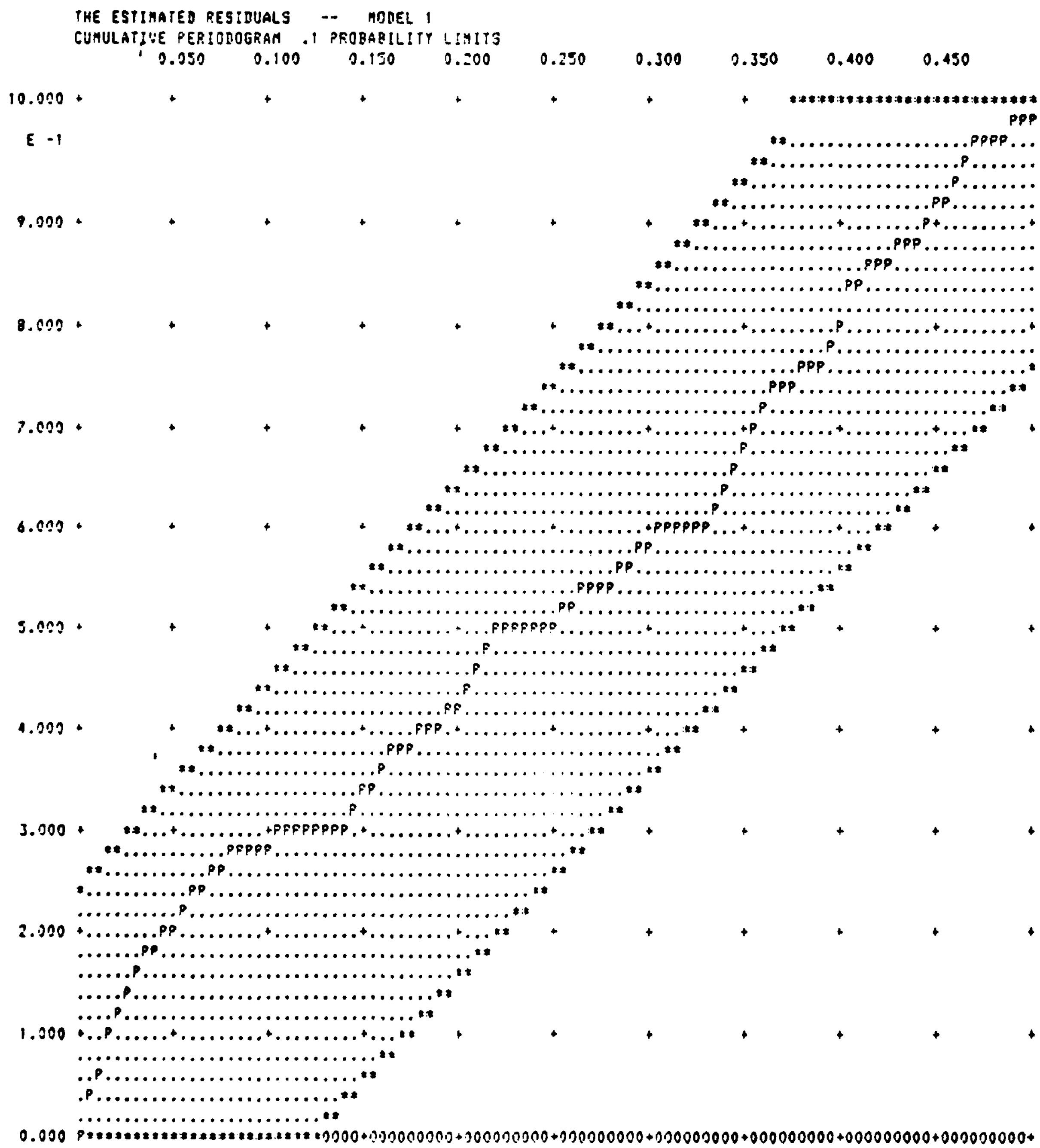


Figure 24. Cummulative Periodogram - Flying Hour Residuals

POWER SPECTRUM
FLYING HOUR RESIDUALS

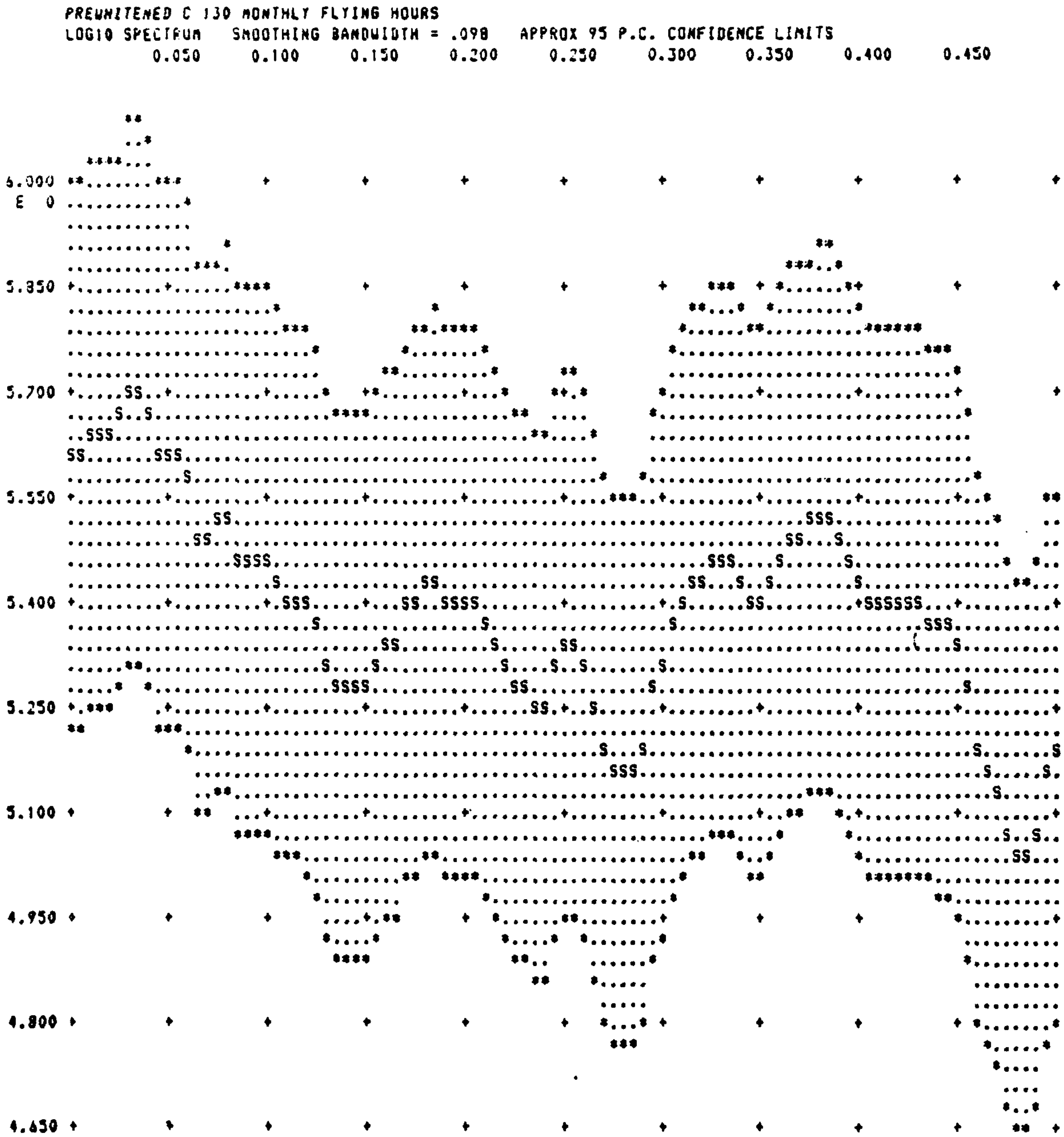


Figure 25. Power Spectrum - Flying Hour Residuals

HISTOGRAM
FLYING HOUR RESIDUALS

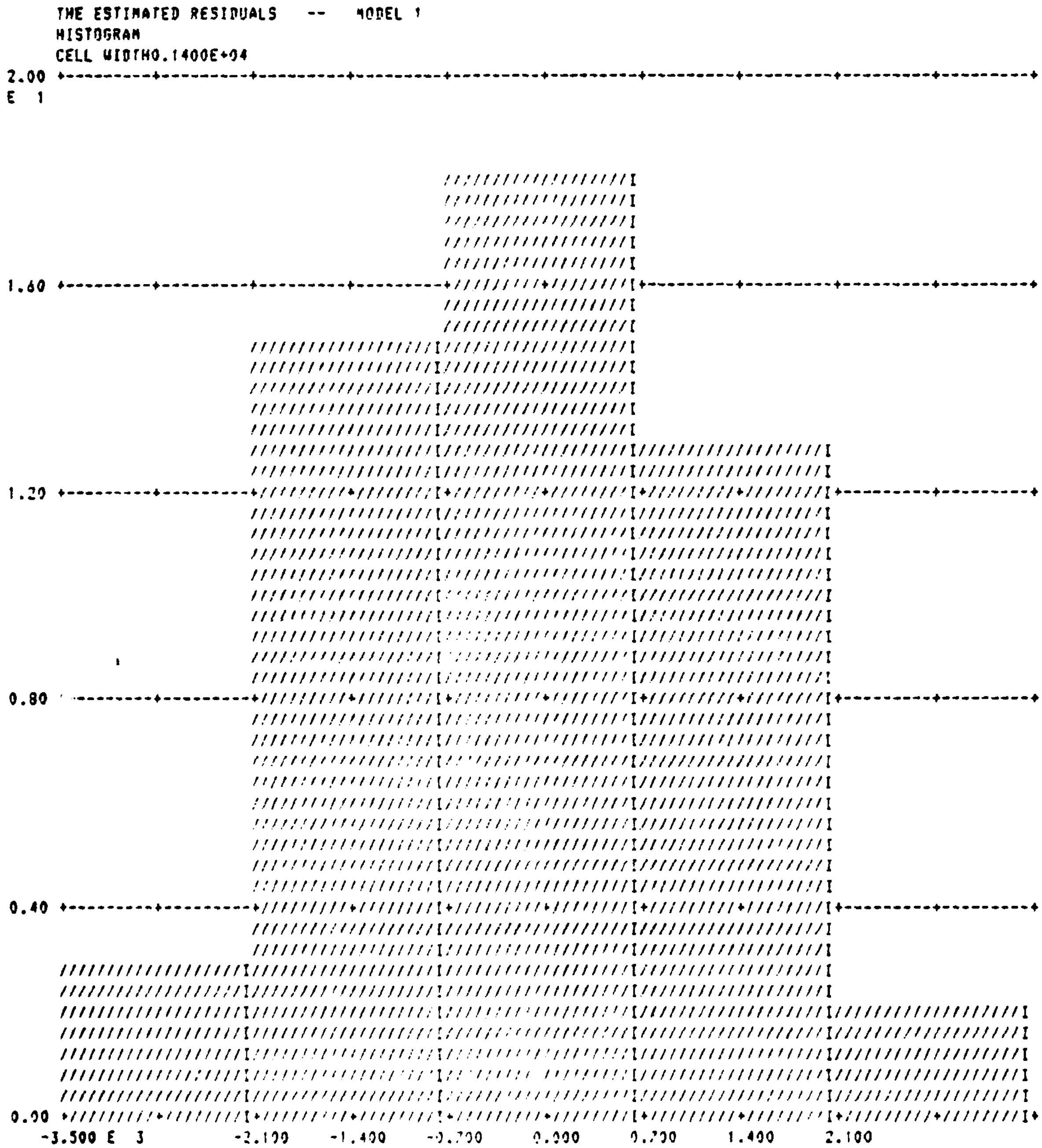


Figure 26. Histogram - Flying Hour Residuals

ESTIMATION RESULTS WITH RECENT DATA

SUMMARY OF MODEL 1

****

DATA - Z = C 130 MONTHLY FLYING HOURS   39 OBSERVATIONS

DIFFERENCING ON Z - NONE

****

PARAMETER NUMBER	PARAMETER TYPE	PARAMETER ORDER	ESTIMATED VALUE	95 PER CENT LOWER LIMIT	95 PER CENT UPPER LIMIT
1	AUTOREGRESSIVE 1	12	0.96513E+00	0.73714E+00	0.11931E+01
2	MEAN	0	0.36813E+05	-.20235E+05	0.93860E+05
3	MOVING AVERAGE 1	1	-.39094E+00	-.80413E+00	0.22250E-01
4	MOVING AVERAGE 2	6	0.79136E+00	0.45295E+00	0.11298E+01
5	MOVING AVERAGE 3	9	-.42815E+00	-.88606E+00	0.29765E-01

****

OTHER INFORMATION AND RESULTS

****

RESIDUAL SUM OF SQUARES	0.41365E+08	22 D.F.	RESIDUAL MEAN SQUARE	0.18802E+07
NUMBER OF RESIDUALS	27		RESIDUAL STANDARD ERROR	0.13712E+04

Figure 27. Estimation Results with Recent Data

PLOT OF CROSS CORRELATION

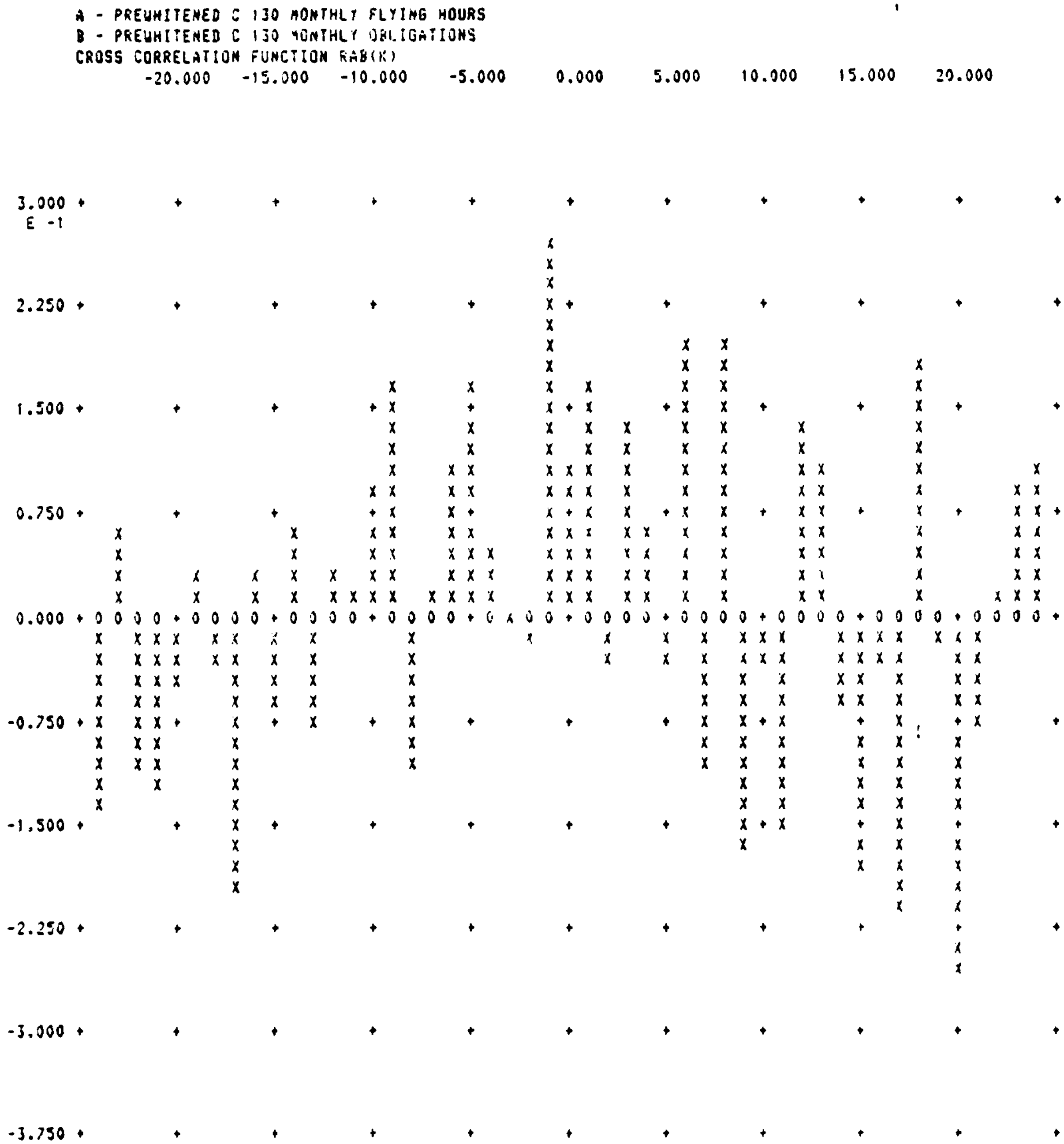


Figure 28. Plot of Cross Correlation

lag -1 is almost significant since it equals .264. If the cross correlation at lag -1 is considered significant, "feedback" would be occurring (i.e., obligations in one month would be causing flying hours in the next month). Since this feedback relationship does not seem intuitively reasonable and since the relationship is very weak, the cross correlation at lag -1 was not considered in determining the form of the transfer function model. The cross correlation at lag -17 is also almost significant (-.216) and more importantly does represent the relationship expected. In computing and buying recoverable spares, the purchase request is prepared to acquire spares to support a future flying hour program (procurement leadtime away). The obligation is recorded in accounting records when the contract for spares is signed by both parties. The flying hour program which the requirement was to support is still production leadtime away. Production leadtime for recoverable spares can be more than 3 years; therefore, 17 months is not an unrealistic average lag or production leadtime. Therefore, obligations could be expected to "lead" flying hours by 17 months. Again, since this relationship at lag -17 is not quite significant (using two standard errors), it was not modeled. In summary, the patterns associated with the significant (.283) cross correlation at lag 20, which indicates that flying hours cause

obligations approximately 20 months later, were examined to determine the form of the transfer function.

The general class of the transfer function is represented by

$$(1 - \delta_1 B - \ldots - \delta_r B^r)\, a_t = (w_0 - w_1 B - \ldots - w_s B^s)\, b_{t-b} + N_t. \qquad (2)$$

The values of r, s, and b are obtained by an examination of patterns and values of the cross correlations. The value of b is the number of cross correlations equal to zero before a significant cross correlation is reached (8:379). The value of b therefore appears to be equal to 20. Since the pattern leading to the spike at lag 20 may start at lag 17 or 18, b values of 17 and 18 will also be tested. The value of r is derived from the pattern of the cross correlation. If r equals 2, the cross correlations are expected to decrease in a damped sine wave pattern after b lags. In addition, the value of r should equal the number of "initial" or "start up" values before the b lag. If r equals 0, a single spike is present at b lags (8:379). Therefore, in this case, r could equal 2 or 0. The value of s is also identified from the patterns used to identify s. The pattern used to identify r begins at lag b + s + 1 (8:379). Therefore, s appears to equal 0. Four transfer function models of the (r,s,b) forms (0,0,20), (2,0,20), (2,0,18) and (2,0,17) were examined. The

AUTOCORRELATION FUNCTION
TRANSFER FUNCTION RESIDUALS

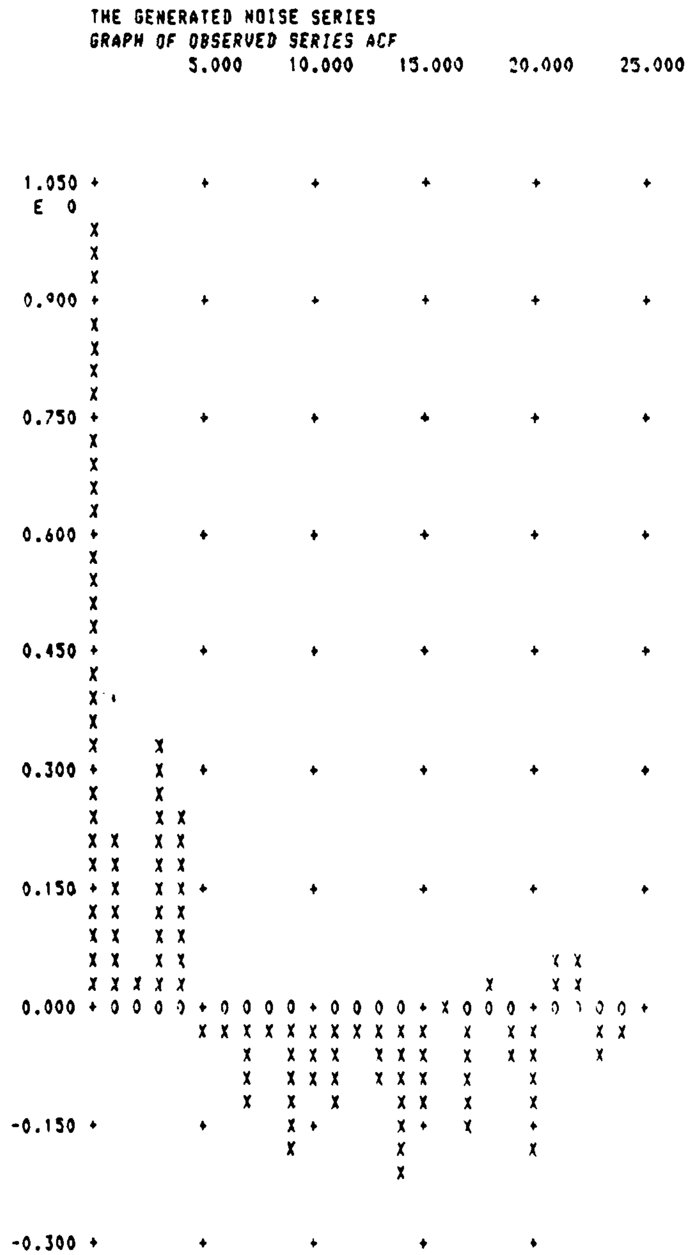


Figure 29. Autocorrelation Function - Transfer Function Residuals

PARTIAL AUTOCORRELATION FUNCTION
TRANSFER FUNCTION RESIDUALS

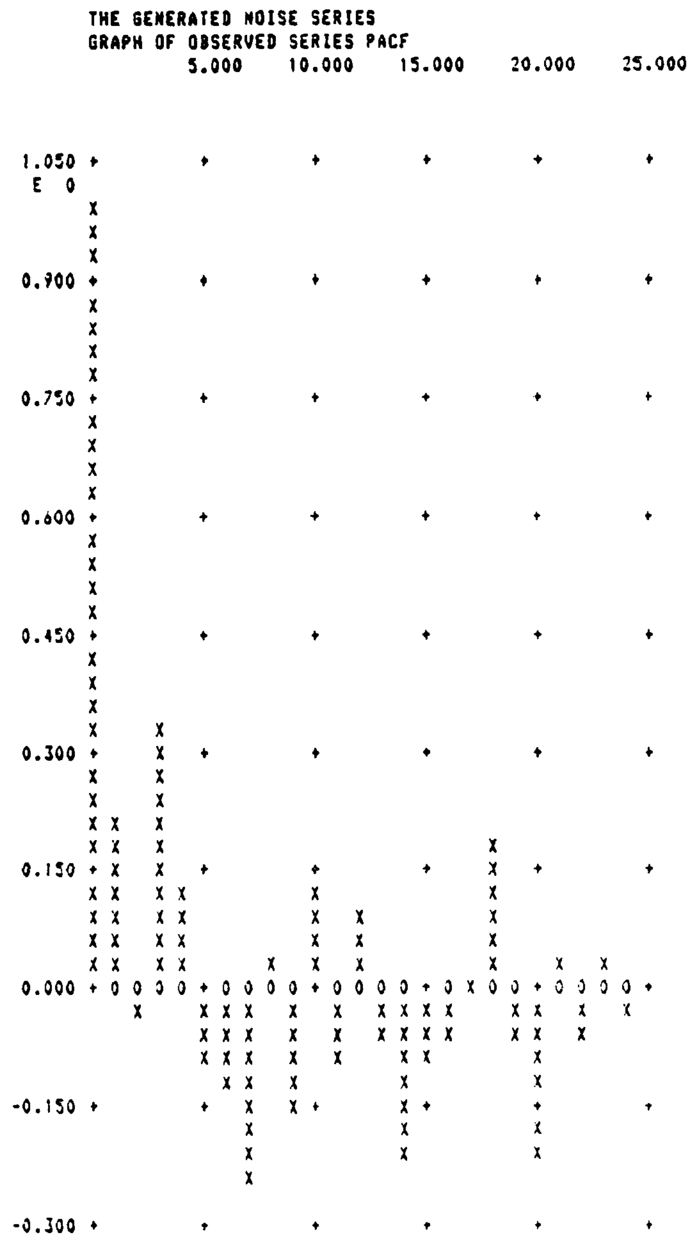


Figure 30. Partial Autocorrelation Function - Transfer Function Residuals

Transfer Function	Chi-Square	Residual Mean Square
0,0,20	10.5 @ 26 d.f.	.34589 E+14
2,0,20	10.7 @ 26 d.f.	.36385 E+14
2,0,18	7.4 @ 26 d.f.	.38755 E+14
2,0,17	9.3 @ 26 d.f.	.32096 E+14

Figure 31. Estimation Results

generated noise series does not exhibit any significant autocorrelation or partial autocorrelation patterns (figures 29 and 30). An attempt to model the slight (.25) ACF and PACF spike at lag 3 did not provide an improved residual mean square. Therefore, noise parameters were deemed as unnecessary.

Transfer Function - Estimation

Parameters for transfer function models (0,0,20), (2,0,20), (2,0,18), and (2,0,17) were estimated.

The results of the estimation (figure 31) are contradictory. An improvement in the "fit" of the model measured by the lower chi-square values was expected to also result in reductions of deviations from the forecast measured by the residual mean square. However, the best fit, a (2,0,18) transfer function model, produces the greatest deviation measured by the residual mean square. These type results could be caused by data which is not truely linear (10). One

important assumption, upon which the entire Box - Jenkins analysis is based, is the assumption of linear relationships. A linear relationship between obligation and flying hours may, in fact, only exist over the middle portion of a weapon systems life cycle. Early in a systems life cycle, relatively high failure rates can be expected which would decrease and then possibly remain linear over a number of years as reliability is improved through modifications. Later in a weapon systems life cycle, as components reach the end of their useful life, the rate of obligations per flying hour could again be expected to increase possibly in a nonlinear fashion. Therefore, the relationship may not be linear early and then late in a weapon systems life cycle. Since the C-130 is an older system the obligation per flying hour rate may be increasing in a rate that is not linear. Although this analysis does not prove the data to be nonlinear, nonlinear indications may be present. Since the model would be used to forecast future obligations, the residual mean square is a stronger indicator of model adequacy. Therefore, a (2,0,17) transfer function model was chosen.

Transfer Function - Diagnostics

The following diagnostics were completed to identify possible model discrepancies and improvements. First, the

plot of residual deviations from the mean (figure 32) was examined. The residual appear to be stationary since the mean and variance do not change over time. In addition, the mean is near zero. A review of the periodogram of the residuals (figure 33) indicates that the residuals are white noise. The chi-square statistic of 9.3 with 26 degrees of freedom also confirm that the residuals are white nose and the model cannot be improved. Plots of the ACF and PACF (figures 34 and 35) reveal no additional inter-series relationships or additional parameters to model the noise series. In summary, the (2,0,17) transfer function model is satisfactory.

Simple Linear Regression

Since both the Hqs AF/ACM and AFLC forecasting methods (POSSEM and ALERT) utilize simple linear regression analysis techniques, the TIMES computer code was used to accomplish a simple linear regression analysis of the data for comparison with the Box Jenkins techniques. The general form of a linear regression equation ($Y_t = B_o + B_1 X_t$, or $Y_t - B_o = B_1 X_t$) can also be stated as a (0,0,0) transfer function model ($Y_t = W_o X_t + e_t$ with Y_t transformed to equal $Y_t - B_o$). An estimate of B_o or the y intercept was computed using the following least squares method formulas (1:403)

DEVIATIONS FROM THE MEAN
TRANSFER FUNCTION RESIDUALS

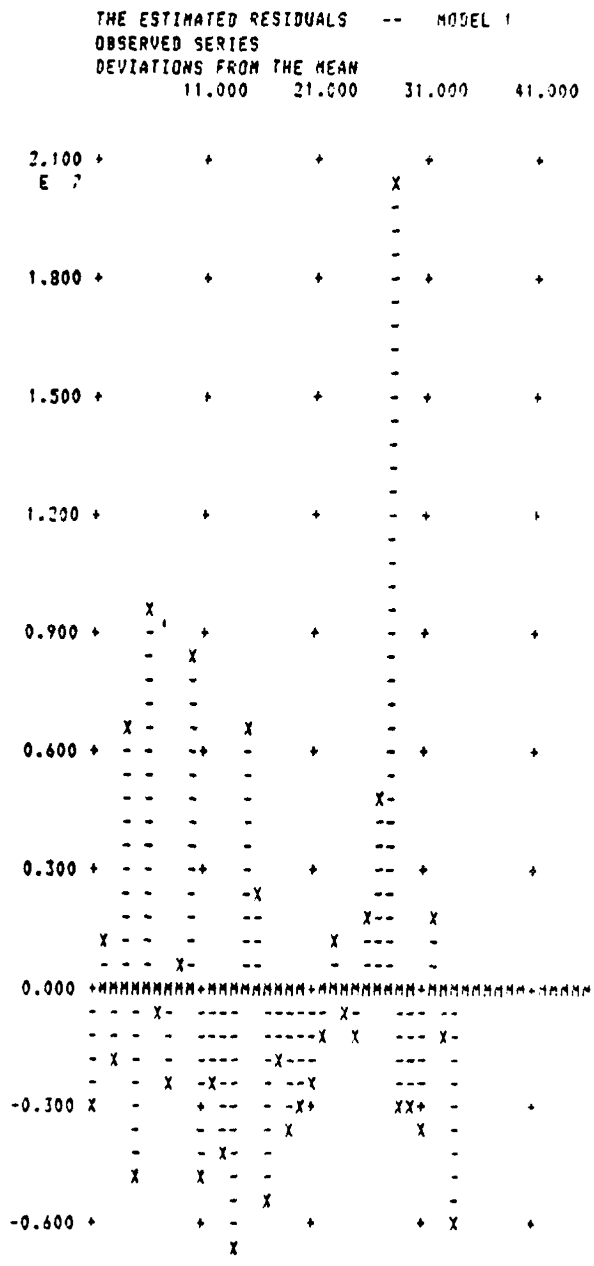


Figure 32. Deviations from the Mean - Transfer Function Residuals

CUMMULATIVE PERIODOGRAM
TRANSFER FUNCTION RESIDUALS

THE ESTIMATED RESIDUALS -- MODEL 1
CUMULATIVE PERIODOGRAM .1 PROBABILITY LIMITS

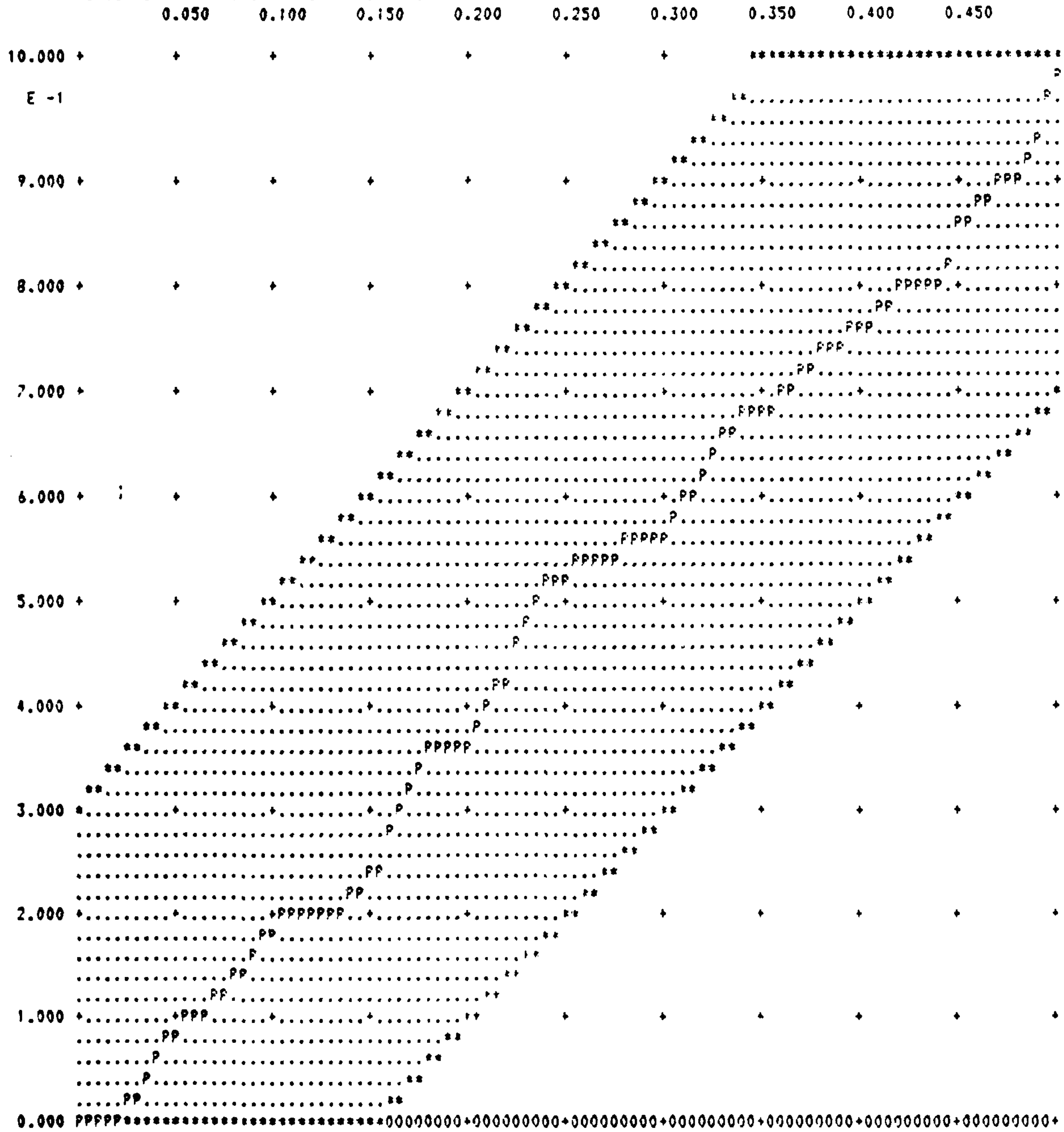


Figure 33. Cummulative Periodogram - Transfer Function Residuals

AUTOCORRELATION FUNCTION
TRANSFER FUNCTION RESIDUALS

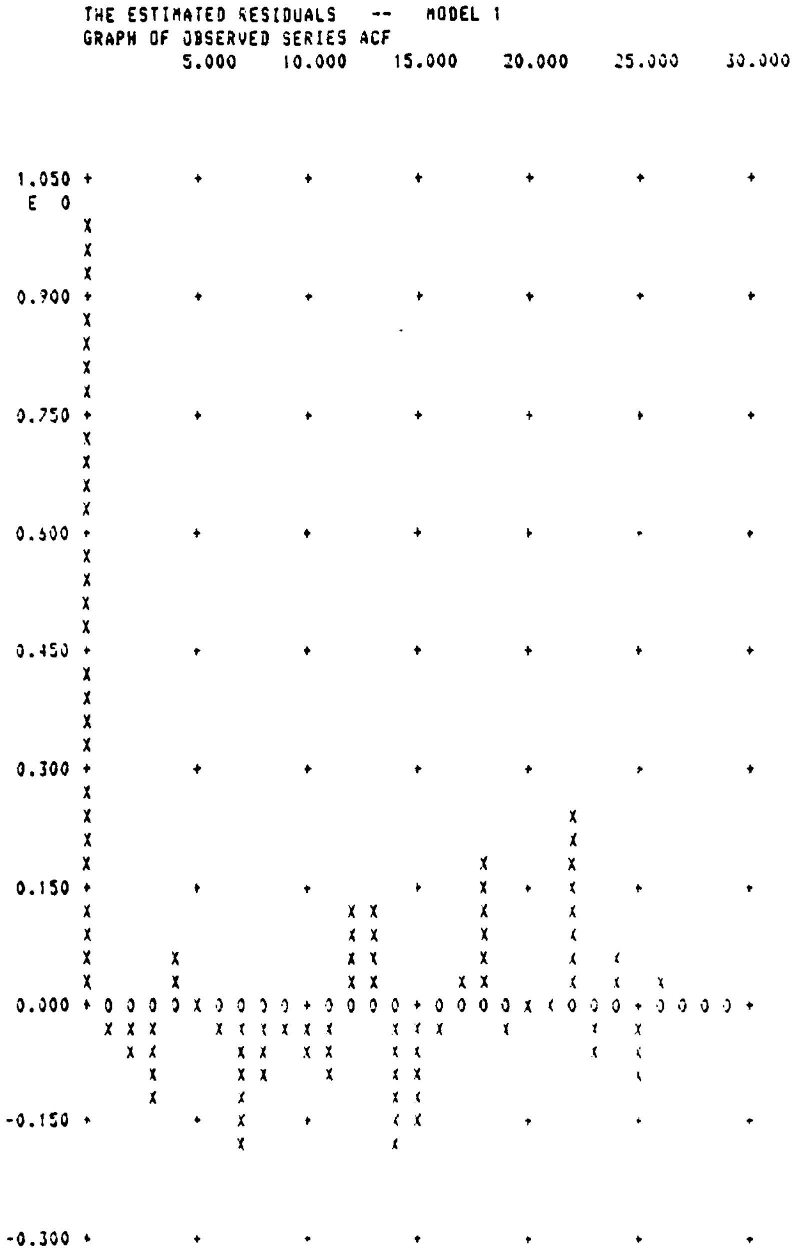


Figure 34. Autocorrelation Function - Transfer Function Residuals

PARTIAL AUTOCORRELATION FUNCTION
TRANSFER FUNCTION RESIDUALS

THE ESTIMATED RESIDUALS -- MODEL 1
GRAPH OF OBSERVED SERIES PACF

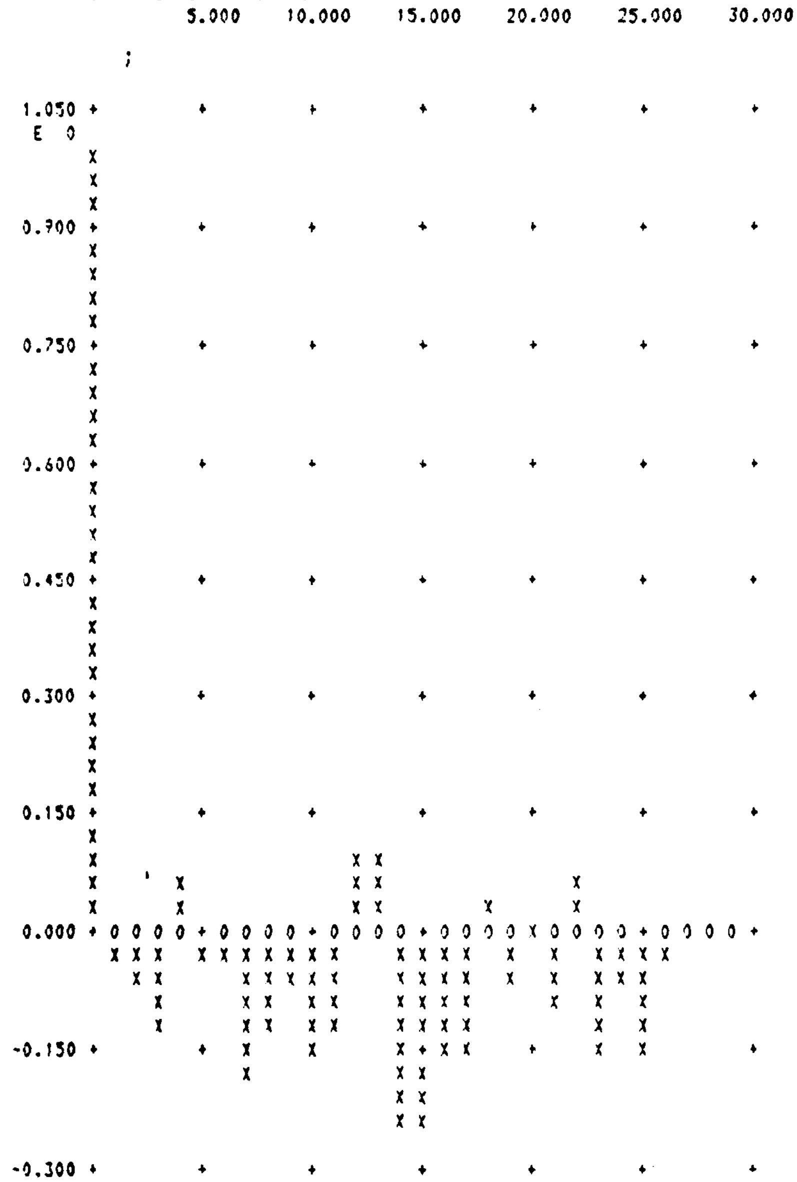


Figure 35. Partial Autocorrelation Function - Transfer Function Residuals

Linear Regression Results	
Chi Square	15.7 @ 31 dif
Residual Mean Square	.25900 E+14

Box Jenkins Results	
Chi-Square	10.9 @ 26 dif
Residual Mean Square	.33461 E+14

Figure 36. Comparison of Linear Regression and Box-Jenkins Methods

$$\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}} \quad (3)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (4)$$

$$SS_{xy} = \sum_{i=1}^{n} x_i y_i - \frac{\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n} \quad (5)$$

$$SS_{xx} = \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n} \quad (6)$$

To obtain comparable statistics the prewhitened obligation data residuals entered in the TIMES code were computed using the formula residuals = Y_t - B_o with B_o = 3,308,217 (computed using the above formula). The original monthly flying hour data were entered unchanged into the TIMES code

as X_t residuals. The results were not conclusive indicating once again that the assumption of linearity may not be correct. A comparison of the Box Jenkins results with the linear regression results is reflected in figure 36. Again, if the relationships were truely linear, the improved "fit" (as measured by the lower chi-square statistic) of the Box Jenkins derived model should produce forecasts with less deviation (a lower residual mean square) (10).

V. CONCLUSIONS AND RECOMMENDATIONS

Findings

The most important finding from the Box - Jenkins analysis of C-130 obligations and flying hours was the identification of relationships within the obligation time series and relationships between the flying hour and obligation time series. Identification and recognition of these relationships resulted first in a slight improvement in the "fit" of the univariate forecast model to the actual obligation data measured by a chi-square improvement from 13.4 with 26 *degrees of freedom* with the original series to 11.7 with 24 degrees of freedom with an ARIMA $(0,0,0)*(0,0,1)_3$ obligation forecast model. Including the relationship between the two times series in a transfer function model resulted in a further improvement of the fit (from 11.7 with 24 degrees of freedom to 9.3 with 26 degrees of freedom). The expected corresponding reduction in the residual mean square did not occur indicating the possibility of non-linear relationships. The linear regression forecasting techniques currently in use do not recognize the possibility of time lag, seasonality, and moving average relationships present between the two time series.

Research Limitation

Box-Jenkins times series analysis requires at least 50 data points. This research effort illustrates the need for even more data. A total of 63 months of flying hour data was used in the analysis. Due to the autoregressive parameter in the flying hour data at lag 12, only 51 residual data points remained for transfer function/noise model analysis. Since the transfer function analysis indicated obligations were related to flying hours 17 months earlier, the residuals from a (2,0,17) transfer function model were limited to only 34 data points. During the various transformations almost half (from 63 to 34) the data points became useless. The small number of remaining residual may contribute to the lack of conclusive results and to the wide range at a 95% confidence interval of parameter values.

Conclusions

This research has attempted to demonstrate that the current linear regression techniques used to forecast obligations can be improved by using a Box Jenkins time series analysis technique to recognize time factors inherent in the data. Although the Box Jenkins techniques produced a model producing a smaller chi-square value indicating a better "fit", the residual mean square, a measure of dispersion of the residuals did not improve as expected. The residual mean

square should decrease as the "fit" is improved if the relationships are in fact linear as has been assumed. The results, therefore, may indicate the presence of some non-linear relationships or may be due to the small number of data points available.

Since the TIMES computer code is available to the Air Force and can be run on any computer capable of compiling FORTRAN ver. 5 (B:72), AFLC has the computational means to identify relationships, not only between obligations and flying hours, but also between obligations and other variables such as fleet value, fleet age, sorties, landings, etc. A serious limitation to more indepth relationship analysis is data availability. Since the obligation data used in this thesis was collected, some of the monthly accounting reports, which were the only data source for this research, have been discarded. Monthly obligation data is only maintained for 3 years. If detailed analysis of time series relationships is to be accomplished in the future, an adequate data base must be developed and maintained beyond 3 years.

Recommendations

As this research was accomplished and conflicting results (decreased chi-square and increase residual mean square) were

obtained, requirements for additional research became apparent.

First, the use of Box-Jenkins techniques to identify relationships between obligations and flying hours with other aircraft should be investigated. This effort should provide a clue to the accuracy of the assumption of linear relationships. Other weapon systems should be selected which are not as old as the C-130 fleet.

Next, relationships between obligations and other variables such as fleet age, fleet value, sorties, landings, etc. should be investigated. Past AFLC research (4) has identified fleet age and fleet value as greater influences on obligations than flying hours (especially with cargo aircraft). Although the C-130 monthly obligation series appears to be random (a chi-square value of 13.4 with 26 degrees of freedom) additional research may identify multiple interrelationships which are in fact very predictable.

The most important recommendation involves the basic cornerstone of the POM forecasting process (forecasting obligations based on expected age, value, flying hours, etc.) and the "Corona Require" (5) findings. To summarize the "Corona Require" finding, in computing POM requirements if

flying hours increase by a certain percent (say 10%) the POM requirement also increases by the same percent. However, in practice if the original requirement totaled 10 assets and 9 assets were already on hand, the POM requirement is the resulting shortfall or the additional asset. A 10% flying hour increase may increase the total requirement by 10% (to 11 assets) but increase the POM (shortfall) requirement by 100% (11 required less 9 on hand) equal 2 short compared to only 1 short before.

The ALERT and POSSEM systems still retain the weakness explained in the "Corona Require" findings. Obligations for spares in any time period are a function of both spares consumption or demands and condemnation rates expected leadtime away and shortages caused by funding constraints in prior periods. In computing BES (more near term) item by item requirements forecasted flying hours are used to develop a forecast of future demands. The total inventory level for a particular item required to support these future demands is then computed considering repair rates, pipeline times, condemnation rates, etc. Finally, the shortage (inventory level required less inventory on hand) is computed. Hence, as noted by the "Corona Require" review a 10% flying hour increase may double the budget requirement.

The recommendation, in short, is to compute long term weapon system inventory level requirements rather than obligations. The POM requirement would then be computed as follows. C-130 Inventory Level Requirement (i.e., FY 1988) less Expected Inventory On Hand less Expected Inventory on Order equals FY 1988 C-130 Obligations Required

A data base needed for this type computation would include the dollar value of monthly world-wide inventory levels by weapon system. This data is available in and could be extracted from the current D041 system. The relationships between monthly inventory levels and flying hours or some other predictive variable (fleet age, fleet value) should be identified using Box-Jenkins techniques.

APPENDIX A. MONTHLY C-130 OBLIGATION DATA

MONTH		MONTHLY OBLIGATION (Non Cumulative)
Jan	Fiscal Year 1980	1,442,684
Feb		497,715
Mar		885,424
Apr		4,605,047
May		600,240
Jun		796,023
Jul		78,339
Aug		1,168,360
Sep		2,100,853
Oct	Fiscal year 1981	104,860
Nov		1,510,373
Dec		1,327,949
Jan		1,252,192
Feb		1,774,317
Mar		372,597
Apr		3,299,785
May		7,079,980
Jun		286,825
Jul		4,498,713
Aug		5,610,453
Sep		10,671,594
Oct	Fiscal Year 1982	360,015
Nov		18,730,663
Dec		3,245,075
Jan		3,538,231
Feb		5,489,859
Mar		15,528,777
Apr		3,098,205
May		1,035,781
Jun		1,614,418
Jul		643,836
Aug		6,409,255
Sep		8,289,767

Oct	Fiscal Year 1983	0
Nov		643,128
Dec		549,191
Jan		288,900
Feb		1,407,428
Mar		2,054,887
Apr		4,960,999
May		4,019,153
Jun		7,316,784
Jul		2,673,377
Aug		5,622,340
Sep		27,852,676
Oct	Fiscal Year 1984	0
Nov		0
Dec		6,707,000[1]
Jan		661,832
Feb		2,630,146
Mar		1,924,201

---

1 The December 1983 report was not available in AFLC/ACB. The December 1983 obligation data (rounded to the nearest thousands) was obtained from AFLC/LO records.

APPENDIX B. MONTHLY C-130 FLYING HOUR DATA

MONTH		MONTHLY FLYING HOURS (Non-Cumulative)
Jan	Fiscal Year 1979	28,004
Feb		26,943
Mar		31,926
Apr		30,606
May		31,418
Jun		32,784
Jul		30,932
Aug		31,842
Sep		29,454
Oct	Fiscal Year 1980	32,353
Nov		29,377
Dec		25,167
Jan		29,257
Feb		28,982
Mar		31,350
Apr		30,470
May		30,544
Jun		29,452
Jul		29,552
Aug		29,182
Sep		30,560
Oct	Fiscal Year 1981	31,199
Nov		27,825
Dec		26,216
Jan		29,639
Feb		29,913
Mar		32,546
Apr		31,290
May		31,546
Jun		31,892
Jul		31,441
Aug		31,729
Sep		30,787

VITA

Thomas G. Lockette was born 30 December 1944 at Eglin AFB, Florida. He graduated from Jacksonville State University at Jacksonville, Alabama with a Bachelor of Science Degree with majors in Math and Accounting in June 1967. After completion of Officer Training School in September 1967, he received a commission in the USAF. He served as budget officer for the 9th Weather Reconnaissance Wing at McClellan AFB, California until returning to civilian life in September 1971. From November 1971 until May 1977, he was employed by various publishing companies leading to a position as comptroller, Oxmoor House Inc., Birmingham, Alabama. In December 1974, he completed the requirements for a Masters of Business Administration Degree from Auburn University. He has been in civil service since 1977 first as an accounting and finance officer and for the last 5 years as an auditor with the Air Force Audit Agency (AFAA). While with the AFAA he completed requirements for and became a Certified Public Accountant in 1981.

Permanent Address: 1301 Boyd Rd.
Xenia, Ohio 45385

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): AFIT/GLM/LSM/84S-38	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: School of Systems & Logistics — 6b. OFFICE SYMBOL (If applicable): AFIT/LS	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code): Air Force Institute of Technology, Wright-Patterson AFB, Ohio 45433	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION — 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State and ZIP Code)	10. SOURCE OF FUNDING NOS.: PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT NO.
11. TITLE (Include Security Classification): See Box 19	

12. PERSONAL AUTHOR(S): Thomas G. Lockette, MBA, GM-13

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
MS Thesis	FROM ______ TO ______	1984 September	96

16. SUPPLEMENTARY NOTATION

Approved for public release: IAW AFR 190-17.
LYNN E. WOLAVER 14 Sept 84
Dean for Research and Professional Development
Air Force Institute of Technology (ATC)
Wright-Patterson AFB OH 45433

17. COSATI CODES		
FIELD	GROUP	SUB. GR.
15	05	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Spares, Spares Requirements, Budget, Forecasting, Spares Forecasting, Times Series Analysis

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Title: A TIME SERIES ANALYSIS OF RECOVERABLE SPARES REQUIREMENTS

Thesis Chairman: Carlos M. Talbott, Major, USAF

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
Carlos M. Talbott, Major, USAF	513-255-5023	AFIT/LSM

DD FORM 1473, 83 APR — EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

The need for a better understanding of relationships between various logistics variables and spares requirements has become evident due to recent understatements of recoverable spares requirements. This thesis used Box-Jenkins time series analysis techniques to identify the linear relationships within and between C-130 spares requirements and flying hours. A methodology will be demonstrated which could then be applied to other variables (fleet value, fleet age, etc.) and other aircraft. A more accurate forecasting technique, than the linear regression technique presently used, should then evolve as more relationships are discovered. The TIMES package on the AFIT Harris computer was used to accomplish the analysis.

The analysis identified relationships within and between the two time series which are not currently considered in the linear regression forecasting model. A model which recognizes these additional relationships (time lags, seasonality, etc.) provides a better "fit" to the historical data, mesured by a chi-square statistic but does not reduce dispersion about the mean, measured by the "residual mean squared". This result is an indication that the relationship may not be linear. Additional data and research is required to confirm that the relationships are non-linear and to develop a model suitable for long range requirements forecasting.

END